



**Universidad Nacional Mayor de San Marcos**

Universidad del Perú. Decana de América

Facultad de Ciencias Biológicas

Escuela Profesional de Genética y Biotecnología

**Determinación de frecuencias alélicas de 20  
marcadores microsatélites (STRs) para la  
identificación humana en una muestra poblacional  
mestiza peruana**

**TESIS**

Para optar el Título Profesional de Biólogo Genetista  
Biotecnólogo

**AUTOR**

Cristian Saul ROBLES MAMANI

**ASESOR**

Margarita VELÁSQUEZ REINOSO

Lima, Perú

2019



Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

<https://creativecommons.org/licenses/by-nc-sa/4.0/>

Usted puede distribuir, remezclar, retocar, y crear a partir del documento original de modo no comercial, siempre y cuando se dé crédito al autor del documento y se licencien las nuevas creaciones bajo las mismas condiciones. No se permite aplicar términos legales o medidas tecnológicas que restrinjan legalmente a otros a hacer cualquier cosa que permita esta licencia.

## Referencia bibliográfica

---

Robles, C. (2019). *Determinación de frecuencias alélicas de 20 marcadores microsatélites (STRs) para la identificación humana en una muestra poblacional mestiza peruana*. Tesis para optar el título profesional de Biólogo Genetista Biotecnólogo. Escuela Profesional de Genética y Biotecnología. Facultad de Ciencias Biológicas, Universidad Nacional Mayor de San Marcos, Lima, Perú.

---

## HOJA DE METADATOS COMPLEMENTARIOS

Código Orcid del autor: <https://orcid.org/0000-0002-9967-7534>

Código Orcid de la asesora: <https://orcid.org/0000-0003-2589-0508>

DNI del autor: 72651104

Grupo de investigación: Genes, cromosomas y genomas de mamíferos.

Institución que financia parcialmente o totalmente la investigación:

Vicerrectorado de Investigación y Posgrado de la Universidad Nacional Mayor de San Marcos, por el financiamiento para Proyectos de Investigación para Grupos de Investigación Año-2018. Proyecto B18101931.

Ubicación geográfica del desarrollo de la tesis:

- Laboratorio de Genética Humana, Facultad de Ciencias Biológicas- Universidad Nacional Mayor de San Marcos. Lima. 12°03'30"S 77°05'00"O.
- Laboratorio de biotecnología y biología molecular- Instituto Nacional de Salud. Chorrillos. -12.182066, -77.016607

Año o rango de años que la investigación abarco: 2018-2019



**UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS**  
(Universidad del Perú, DECANA DE AMÉRICA)

**FACULTAD DE CIENCIAS BIOLÓGICAS**

**ACTA DE SESIÓN PARA OPTAR AL TÍTULO PROFESIONAL DE  
BIÓLOGO GENETISTA BIOTECNÓLOGO  
(MODALIDAD: SUSTENTACIÓN DE TESIS)**

Siendo las 10:16 horas del 9 de octubre de 2019, en el Salón de Grados de la Facultad de Ciencias Biológicas y en presencia del jurado formado por los profesores que suscriben, se dio inicio a la sesión para optar al **Título Profesional de Biólogo Genetista Biotecnólogo** de **CRISTIAN SAUL ROBLES MAMANI**.

Luego de dar lectura y conformidad al expediente N° 012-EPGB-2019, el titulando expuso su tesis: **“DETERMINACIÓN DE FRECUENCIAS ALÉLICAS DE 20 MARCADORES MICROSATÉLITES (STRs) PARA LA IDENTIFICACIÓN HUMANA EN UNA MUESTRA POBLACIONAL MESTIZA PERUANA”** y el Jurado efectuó las preguntas del caso calificando la exposición con la nota 19, calificativo: Aprobado con máximos honores.

Finalmente, el expediente será enviado a la Escuela Profesional de Genética y Biotecnología, y al Consejo de Facultad para que se apruebe otorgar el **Título Profesional de Biólogo Genetista Biotecnólogo** a **CRISTIAN SAUL ROBLES MAMANI** y se eleve lo actuado al Rectorado para conferir el respectivo título, conforme a ley.

Siendo las 12:00 horas se levantó la sesión.

Ciudad Universitaria, 9 de octubre de 2019.

Mg. OLGA BRACAMONTE GUEVARA  
(PRESIDENTA)

Blga. MARGARITA VELÁSQUEZ REINOSO  
(ASESORA)

Blgo. ALBERTO LOPEZ SOTOMAYOR  
(MIEMBRO)

Blgo. JORGE ZEBALLOS ALVA  
(MIEMBRO)

## **DEDICATORIA**

*A mis padres, a mis hermanos y a mi querido perrito.*

## **AGRADECIMIENTOS**

A mis padres, por ser pilar fundamental en esta empresa, por permitirme lograr este anhelado sueño lleno de dificultades y retos.

A mi asesora y profesora, Margarita Velásquez Reinoso, por aceptarme en formar parte de tan ilustre Laboratorio de Genética Humana, por la guía, paciencia y dedicación en la culminación de este designio. Asimismo al Dr. Jaime Descailleaux por sus acertados consejos.

Al Dr. Carlos Neyra Rivera, por sus sapientes consejos, por confiar desde un principio en mi trabajo, por orientarme e inculcarme los conocimientos básicos en genética forense.

Al Laboratorio de Biotecnología y Biología Molecular del INS, por permitirme hacer uso de sus equipos, los cuales fueron piezas invaluable en esta investigación.

A los profesores Alberto López, Jorge Zeballos y Olga Bracamonte, por las sugerencias asignadas y por dedicar su tiempo a la revisión de este escrito.

A la Mg. Claudia Barletta, al Dr. Jorge Ramírez y al profesor Jesús Córdova, por los consejos brindados en la realización de este trabajo.

A María Solís, por su tiempo y sugerencias académicas.

A mis colegas y mentores, July, Ángelo y Mari, quienes me dejaron formar parte de su equipo y formaron en mí los primeros cimientos de la biología de la reproducción humana, por permitirme conocer la realidad de los hospitales y la condición del biólogo.

A mis amigos Guido, Josmar, Renzo, Manuel y Otalora, por las innumerables aventuras vividas durante la travesía del pregrado.

Al grupo de estudio de Neurobiología y Fitoquímica (GNFIT), con los cuales aprendí a dar los primeros pasos en la investigación científica y a trabajar en equipo.

A todas las personas que directa o indirectamente han contribuido en la ejecución de este proyecto.

Al Vicerrectorado de Investigación y Posgrado de la Universidad Nacional Mayor de San Marcos, por el financiamiento para Proyectos de Investigación para Grupos de Investigación Año-2018. Proyecto B18101931.

## ÍNDICE GENERAL

ÍNDICE GENERAL .....	i
ÍNDICE DE FIGURAS .....	vi
ÍNDICE DE TABLAS .....	viii
RESUMEN .....	ix
ABSTRACT .....	x
I. INTRODUCCIÓN.....	1
II. MARCO TEÓRICO.....	5
2.1. Características de los STRs .....	7
2.2. Clasificación de los STRs .....	8
2.3. Nomenclatura de los STRs.....	8
2.4. Electroforesis capilar.....	15
III. HIPÓTESIS Y OBJETIVOS .....	19
3.1. Hipótesis.....	19
3.1.1. Hipótesis Estadísticas .....	19
3.2. Objetivo General .....	19
3.2.1. Objetivos Específicos.....	19
IV. MATERIALES Y MÉTODOS.....	20
4.1. Población de Estudio .....	20
4.2. Materiales.....	21
4.2.1. Reactivos.....	21
4.2.2. Consumibles.....	21



4.2.3. Equipos.....	22
4.3. Métodos.....	22
4.3.1. Toma de muestra .....	22
4.3.2. Amplificación de ADN para Electroforesis Capilar .....	23
4.3.3. Electroforesis Capilar .....	24
4.3.4. Identificación de los alelos .....	25
4.3.5. Métodos de Anàlisis e Interpretaciòn de Resultados .....	25
4.3.5.1. Parámetros de interés poblacional.....	25
4.3.5.1.1. Estimación de las frecuencias alélicas.....	25
4.3.5.1.2. Índice de Heterocigosidad.....	25
4.3.5.1.3. Equilibrio de Hardy- weinberg.....	26
4.3.5.1.4. Frecuencias mínimas.....	28
4.3.5.1.5. Desequilibrio de Ligamiento.....	29
4.3.5.2. Parámetros de Interes Forense.....	30
4.3.5.2.1. Índice de Contenido Polimórfico (PIC).....	30
4.3.5.2.2. Probabilidad de Coincidencia (PC).....	30
4.3.5.2.3. Poder de Discriminación (PD).....	31
4.3.5.2.4. Poder de Exclusión (PE).....	31
4.3.5.3. Diferenciación Poblacional.....	32
4.3.5.3.1. Índice de Fijación.....	33
<b>V. RESULTADOS.....</b>	<b>34</b>
5.1. Marcadores STRs analizados en la muestra poblacional mestiza peruana.....	34
5.1.1. Marcador D3S1358 .....	34
5.1.2. Marcador vWA.....	34
5.1.3. Marcador D16S539 .....	34

5.1.4. Marcador CSF1PO.....	34
5.1.5. Marcador TPOX .....	34
5.1.6. Marcador D8S1179 .....	35
5.1.7. Marcador D21S11 .....	35
5.1.8. Marcador D18S51 .....	35
5.1.9. Marcador D2S441 .....	35
5.1.10. Marcador D19S433 .....	35
5.1.11. Marcador TH01 .....	36
5.1.12. Marcador FGA.....	36
5.1.13. Marcador D22S1045 .....	36
5.1.14. Marcador D5S818 .....	36
5.1.15. Marcador D13S317 .....	36
5.1.16. Marcador D7S820 .....	37
5.1.17. Marcador D10S1248 .....	37
5.1.18. Marcador D1S1656 .....	37
5.1.19. Marcador D12S391 .....	37
5.1.20. Marcador D2S1338 .....	37
5.2. Análisis de los parámetros poblacionales.....	41
5.3. Variabilidad de las frecuencias alélicas de la muestra poblacional mestiza peruana e hispanoamericana.....	44
5.3.1. Marcador D3S1358.....	44
5.3.2. Marcador vWA.....	45
5.3.3. Marcador D16S539.....	46
5.3.4. Marcador CSF1PO.....	47
5.3.5. Marcador TPOX.....	48
5.3.6. Marcador D8S1179.....	49

5.3.7. Marcador D21S11 .....	50
5.3.8. Marcador D18S51 .....	51
5.3.9. Marcador D2S441 .....	52
5.3.10. Marcador D19S433 .....	53
5.3.11. Marcador TH01 .....	54
5.3.12. Marcador FGA .....	55
5.3.13. Marcador D22S1045 .....	56
5.3.14. Marcador D5S818 .....	57
5.3.15. Marcador D13S317 .....	58
5.3.16. Marcador D7S820 .....	59
5.3.17. Marcador D10S1248 .....	60
5.3.18. Marcador D1S1656 .....	61
5.3.19. Marcador D12S391 .....	62
5.3.20. Marcador D2S1338 .....	63
5.4. Análisis de parámetros forenses .....	64
5.5. Análisis de diferenciación poblacional .....	66
5.5.1. Análisis de diferenciación poblacional con 13 marcadores STRs entre 23 poblaciones .....	66
5.5.2. Análisis de diferenciación poblacional con 20 marcadores STRs entre 6 poblaciones .....	73
<b>VI. DISCUSIÓN .....</b>	<b>77</b>
6.1. Parámetros poblacionales .....	77
6.1.1. Tamaño y característica de la población .....	77
6.1.2. Análisis de los parámetros poblacionales .....	80
6.2. Parámetros forenses .....	87

6.3. Análisis de diferenciación poblacional.....	89
<b>VII. CONCLUSIONES.....</b>	<b>92</b>
<b>VIII. RECOMENDACIONES .....</b>	<b>93</b>
<b>IX. REFERENCIAS BIBLIOGRÁFICAS.....</b>	<b>94</b>
<b>X. ANEXOS.....</b>	<b>101</b>
ANEXO 1. Consentimiento Informado y Ficha de Registro .....	101
ANEXO 2. Constancia de Aprobación CIEI-2018-015.....	102
ANEXO 3. Lugar de nacimiento de los 200 voluntarios que accedieron a participar en el estudio.....	103
ANEXO 4. Voluntarios nacidos en Limay/o Callao cuyos ancestros pertenecen a un mismo departamento, diferente a Lima y/o Callao.....	104
ANEXO 5. Lugar de nacimiento de los 200 voluntarios participantes en este estudio, así como el lugar de nacimiento de los padres y abuelos paternos y maternos.....	105

## ÍNDICE DE FIGURAS

<b>Figura 1.</b> Estructura de una Repetición Corta en Tándem (STR).....	6
<b>Figura 2.</b> Ubicación cromosómica de los 20 marcadores genéticos microsatélites del CODIS y el marcador genético del sexo, amelogenina. ....	7
<b>Figura 3.</b> Esquema de las posiciones del cebador de PCR para la amplificación de un marcador STR de ADN .....	16
<b>Figura 4.</b> Ilustración esquemática de la separación y detección de alelos STRs con un analizador genético ABI Prism 310.....	18
<b>Figura 5.</b> Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR D3S1358.....	44
<b>Figura 6.</b> Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR vWA. ....	45
<b>Figura 7.</b> Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR D16S539.....	46
<b>Figura 8.</b> Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR CSF1PO. ....	47
<b>Figura 9.</b> Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR TPOX. ....	48
<b>Figura 10.</b> Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR D8S1179.....	49
<b>Figura 11.</b> Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR D21S11.....	50
<b>Figura 12.</b> Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR D18S51.....	51
<b>Figura 13.</b> Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR D2S441.....	52
<b>Figura 14.</b> Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR D19S433.....	53

<b>Figura 15.</b> Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR TH01 .....	54
<b>Figura 16.</b> Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR FGA. ....	55
<b>Figura 17.</b> Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR D22S1045.....	56
<b>Figura 18.</b> Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR D5S818.....	57
<b>Figura 19.</b> Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR D13S317.....	58
<b>Figura 20.</b> Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR D7S820.....	59
<b>Figura 21.</b> Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR D10S1248.....	60
<b>Figura 22.</b> Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR D1S1656.....	61
<b>Figura 23.</b> Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR D12S391.....	62
<b>Figura 24.</b> Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR D2S1338.....	63
<b>Figura 25.</b> Heatmap para 13 marcadores STRs. ....	71
<b>Figura 26.</b> Árbol UPGMA para 13 marcadores STRs. ....	72
<b>Figura 27.</b> Heatmap para 20 marcadores STRs. ....	75
<b>Figura 28.</b> Árbol UPGMA para 20 marcadores STRs .....	76

## ÍNDICE DE TABLAS

<b>Tabla 1.</b> Cuadro resumen de los 20 marcadores STRs autosómicos. ....	14
<b>Tabla 2.</b> Frecuencias alélicas y parámetros de interés poblacional .....	38
<b>Tabla 3.</b> Valores p para el análisis de Desequilibrio de Ligamiento.....	43
<b>Tabla 4.</b> Parámetros de interés forense. ....	65
<b>Tabla 5.</b> Parámetros de interés forense combinado. ....	65
<b>Tabla 6.</b> Poblaciones utilizadas para el análisis de 13 marcadores STRs, enumeradas del 1 al 23.....	69
<b>Tabla 7.</b> Valores $F_{st}$ promedio obtenidos en el análisis entre pares de poblaciones para los 13 marcadores STRs.....	70
<b>Tabla 8.</b> Poblaciones, enumeradas del 1 al 6, utilizadas para el análisis de 20 marcadores STRs. ....	75
<b>Tabla 9.</b> Valores $F_{st}$ promedio obtenidos en el análisis entre pares de poblaciones para los 20 marcadores STRs. ....	75

## RESUMEN

Se determinó la distribución de las frecuencias alélicas de 20 marcadores STRs autosómicos (CSF1PO, FGA, TH01, TPOX, VWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, D1S1656, D2S441, D2S1338, D10S1248, D12S391, D19S433, D22S1045) en una muestra poblacional de 200 individuos no emparentados de mestizos peruanos provenientes de diferentes departamentos del Perú, siendo la gran mayoría provenientes de la provincia de Lima. La caracterización se realizó mediante electroforesis capilar usando el Kit de amplificación de PCR VeriFiler Express. Los datos genéticos se analizaron utilizando los programas Arlequín 3.5.2.2., PowerStats V12, PHYLIP 3.695 y MEGA 6.06; con fines de comparación poblacional, se utilizó información de poblaciones obtenidas de la literatura revisada. Todos los locis analizados estuvieron en equilibrio de Hardy-Weinberg, luego de aplicar la corrección de Bonferroni al marcador D19S433 ( $p=0.0382$ ). Asimismo, la prueba de desequilibrio de ligamiento descartó asociación entre todos los pares de locis luego de aplicar la corrección de Bonferroni. Se determinó la heterocigosidad observada y esperada, así como la frecuencia mínima. La muestra poblacional analizada registró alelos que no han sido descritos en la base de datos del NIST, siendo estos el 12.3 del marcador CSF1PO, 8.2 del marcador TPOX, 9.1 del marcador D2S441, 11.2 del marcador D19S433, 9.2 del marcador D13S317 y 14, 19.1, 20.3 del marcador D12S391. Asimismo, se encontró alelos que no han sido registrados en la población hispanoamericana: los alelos 11 y 12 para el marcador D3S1358, el alelo 13 para el marcador vWA, el alelo 12.3 para el marcador CSF1PO, los alelos 8.2 y 13 para el marcador TPOX, los alelos 34.2 y 35 para el marcador D21S11, el alelo 25 para el marcador D18S51, el alelo 9.1 para el marcador D2S441, los alelos 11.2, 14.1 y 17 para el marcador D19S433, el alelo 19 para el marcador D22S1045, los alelos 9.2 y 17 para el marcador D13S317, el alelo 14 para el marcador D7S820, el alelo 9 para el marcador D10S1248, los alelos 14 y 20.3 para el marcador D12S391. Los parámetros forenses estimados fueron: poder de discriminación (PD), poder de exclusión (PE), índice de contenido polimórfico (PIC) y la probabilidad de coincidencia (PC). El PD y PE combinado para los 20 marcadores microsatélites fue 0.999999999, 0.999999672, respectivamente. La comparación de diferenciación genética a través del estadístico  $F_{st}$  (basada en la distancia genética de Reynolds) con grupos poblacionales de Estados Unidos y otras poblaciones mundiales reveló subdivisión genética entre la muestra poblacional analizada y la hispanoamericana; esta afirmación se corroboró mediante el árbol UPGMA.

**Palabras claves:** microsatélites, frecuencia alélica, electroforesis capilar, parámetros forenses, diferenciación genética.



## ABSTRACT

The distribution of allelic frequencies of 20 autosomal STRs markers (CSF1PO, FGA, TH01, TPOX, VWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, D1S1656, D2S441, D2S1338, D10S1248, D12S391, D19S433, D22S1045) in a population sample of 200 unrelated individuals of Peruvian mestizos from different departments of Peru was determined, the vast majority coming from the province of Lima. The characterization was performed by capillary electrophoresis using the VeriFiler Express PCR amplification kit. Genetic data were analyzed using the Arlequin 3.5.2.2., PowerStats V12, PHYLIP 3.695 and MEGA 6.06 programs; for the purpose of population comparison, information on populations obtained from the literature reviewed was used. All loci analyzed were in Hardy-Weinberg equilibrium, after applying Bonferroni correction to marker D19S433 ( $p=0.0382$ ). Likewise, the linkage disequilibrium test discarded association between all pairs of loci after applying Bonferroni correction. The observed and expected heterozygosity was determined, as well as the minimum frequency. The population sample analyzed recorded alleles that have not been described in the NIST database, being these 12.3 of the marker CSF1PO, 8.2 of the marker TPOX, 9.1 of the marker D2S441, 11.2 of the marker D19S433, 9.2 of the marker D13S317 and 14, 19.1, 20.3 of marker D12S391. Likewise, alleles were found that have not been registered in the Hispano American population: alleles 11 and 12 for marker D3S1358, allele 13 for marker vWA, allele 12.3 for marker CSF1PO, alleles 8.2 and 13 for marker TPOX , alleles 34.2 and 35 for marker D21S11, allele 25 for marker D18S51, allele 9.1 for marker D2S441, alleles 11.2, 14.1 and 17 for marker D19S433, allele 19 for marker D22S1045, alleles 9.2 and 17 for marker D13S317, allele 14 for marker D7S820, allele 9 for marker D10S1248, alleles 14 and 20.3 for marker D12S391.

The forensic parameters were estimated: power of discrimination (PD), power of exclusion (PE), polymorphism information content (PIC) and the probability of coincidence (PC). The combined PD and PE for the 20 microsatellite markers was 0.999999999, 0.999999672, respectively. The comparison of genetic differentiation through the  $F_{st}$  statistic (based on the genetic distance of Reynolds) with population groups in the United States and other world populations revealed genetic subdivision between the sample analyzed and the Hispanic American population; This affirmation was corroborated by the UPGMA tree.

**Keywords:** microsatellites, allelic frequency, capillary electrophoresis, forensic parameters, genetic differentiation.

## 1. INTRODUCCIÓN

La ciencia forense, conocida en algunos países como medicina legal, es una especialidad de la Medicina que tiene como objetivo apoyar a los jueces y jurados a resolver asuntos legales, no solo en el derecho penal sino también en el derecho civil. El campo de la ciencia forense tiene una gran amplitud, cruzando las fronteras entre biología, química, física y matemáticas, e incluye disciplinas tan variadas como botánica y balística, así como el análisis de huellas dactilares, impresiones auditivas, sonido grabado y escritura a mano (Jobling *et al.*, 2004). Durante los últimos 20 años, sin embargo, una herramienta biológica particular ha revolucionado las investigaciones forenses: el análisis del Acido Desoxirribonucleico (ADN). Como todos los seres vivos contienen ADN, y éste exhibe variabilidad tanto dentro como entre las especies, en tal sentido, cualquier material biológico asociado con un caso legal lleva información sobre su origen (Jobling *et al.*, 2004). Es así que la Genética Forense consiste en la aplicación del análisis genético de la diversidad humana para la resolución de ciertos problemas judiciales (Carracedo, 1996). Por lo tanto, la labor del genetista forense es realizar el análisis del material recuperado de la escena del crimen, de las pruebas de paternidad y la identificación de restos humanos (Goodwin *et al.*, 2011).

La moderna tecnología de caracterización de ADN ha abierto nuevas posibilidades de realizar pruebas de identificación humana. La identificación individual es imprescindible en un número de situaciones incluyendo la determinación de autores de violentos crímenes tales como asesinato y agresión sexual, y la identificación de restos de personas desaparecidas o víctimas de desastres masivos (El-Alfy *et al.*, 2012; Ziętkiewicz *et al.*, 2012). Debido a que las condiciones ambientales de un desastre en masa frecuentemente resultan en severa fragmentación, descomposición, y mezcla

de los restos de las víctimas, la identificación tradicional basada en características antropológicas y físicas de las víctimas es frecuentemente poco concluyente. Esta es la razón por la que el perfil de ADN se ha convertido en un estándar de oro (*gold standard*) para la identificación de individuos (Ziętkiewicz *et al.*, 2012).

La mayoría de los sistemas de identificación forense se basan en loci genéticos con secuencias de ADN repetitivas en tándem. Se pueden encontrar secuencias repetidas en el ADN satelital, pero desde el punto de vista forense las regiones de ADN repetitiva mucho más cortas que el ADN satelital son considerablemente más interesantes. Estas regiones se pueden clasificar en minisatélites y microsatélites. Los minisatélites, también conocidos como número variable de repeticiones en tándem o VNTR (*Variable Number of Tandem Repeats*), se componen de motivos de secuencia que varían entre aproximadamente 15 y 50 pb de longitud, reiterados en tándem para alcanzar una longitud total de 500 pb a 20 kpb (Carracedo *et al.*, 2008). Los microsatélites o STRs (*Short Tandem Repeat*) son mucho más cortos y la unidad de repetición varía de 2 a 6 pb (Ruitberg *et al.*, 2001; Carracedo *et al.*, 2008; El-Alfy *et al.*, 2012; Willems *et al.*, 2014) para una longitud total entre 50 pb y 500 pb. Otra fuente de variación ha atraído el interés de los genetistas forenses: los polimorfismos de un solo nucleótido o SNP (*Single Nucleotide Polymorphism*). Estos representan el tipo más simple de polimorfismos (una variación simple en una sola base) y generalmente son marcadores bialélicos (Carracedo *et al.*, 2008).

Los microsatélites son extraordinariamente adecuados para el análisis de muestras forenses que contienen cantidades limitadas o degradadas de ADN (Butler *et al.*, 2004; Tamaki *et al.*, 2005), debido a que se requiere solamente

bajas cantidades de ADN molde (10-100 ng) (Hameed *et al.*, 2014), pudiendo ser fácilmente amplificados con la Reacción en Cadena de la Polimerasa (PCR) (Ruitberg *et al.*, 2001). Los STRs se caracterizan por ser codominantes, altamente polimórficos y casi selectivamente neutros (Fan *et al.*, 2007). También pueden mostrar polimorfismos sustanciales, siendo abundantes en todo el genoma humano (Tamaki *et al.*, 2005; Willems *et al.*, 2014), encontrándose dentro de las regiones codificantes de genes, intrones y en las secuencias no genéticas (Hameed *et al.*, 2014). El número de repeticiones en los marcadores STRs es altamente variable entre individuos, haciendo a estos marcadores efectivos para propósitos de identificación humana, adicionalmente el pequeño tamaño de los alelos STR comparados a los minisatélites hacen a los marcadores STR mejores candidatos para el uso en aplicaciones forenses (El-Alfy *et al.*, 2012). Por medio de éstos se logra elaborar los perfiles genéticos, los cuales son una representación numérica que toma como base la estructura del ADN que se caracteriza por no contener información referente a las características físicas, psicológicas, o patológicas de una persona (Alleyne, 2009).

Para que el proceso de identificación sea más consistente, diferentes países han desarrollado tablas de frecuencias alélicas con su propia población a fin de que los valores de identificación obtenidos sean los más confiables posibles (Budowle *et al.*, 2001; Paredes *et al.*, 2003).

Por lo expuesto, la problemática en el Perú surge debido a que la mayoría de los marcadores empleados para la identificación de la población peruana están sometidos a frecuencias alélicas de bases de datos de muestras poblacionales de inmigrantes hispanoamericanos que residen en los Estados Unidos de Norteamérica (Tito, 2003). Por ende la estimación de las frecuencias alélicas

para estos marcadores microsatélites no representan fielmente la composición genética de la población peruana, por lo que es necesario determinar la variabilidad y las frecuencias alélicas de los marcadores microsatélites usados con propósitos de identificación humana, así como para pruebas de paternidad.

Bajo esta perspectiva, esta muestra mestiza de 200 individuos no relacionados de ambos sexos tomada en la ciudad de Lima-Perú, se caracteriza por ser un grupo poblacional de descendientes de nativos americanos, españoles y en menor grado africanos y asiáticos (Talledo *et al.*, 2010), la cual permitirá conocer la variabilidad y las frecuencias alélicas de 20 marcadores STRs usados para la identificación de personas así como para pruebas de paternidad; adicionalmente, se incrementara la base de datos de ADN forense peruana. Asimismo, con los resultados obtenidos se analiza los parámetros genéticos poblacionales, la eficiencia forense, la estructura poblacional, y se compara las frecuencias alélicas obtenidas con las registradas en otras poblaciones a nivel mundial.

## 2. MARCO TEÓRICO

Con los avances en las técnicas de biología molecular, ahora es posible analizar cualquier región del ADN de los 3,2 billones de bases que conforman el genoma humano (Goodwin et al., 2011).

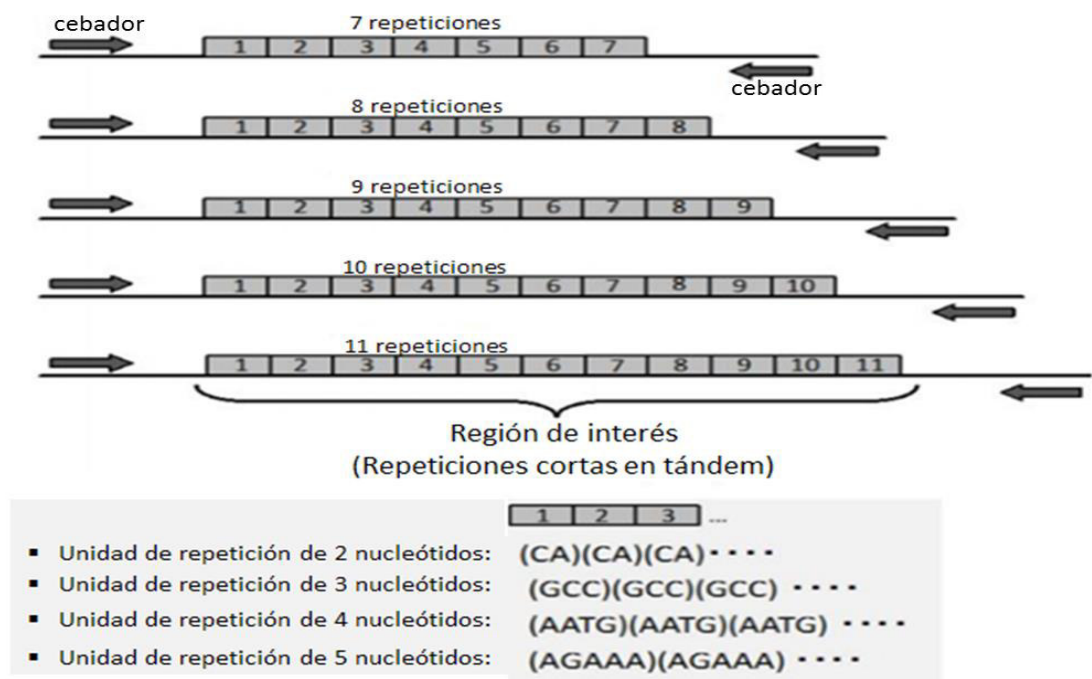
La tipificación de ADN forense tiene un importante impacto en la sociedad porque proporciona una confiable evidencia para condenar a los delincuentes o para exonerar a los sospechosos inocentes. Otro campo de aplicación de la tipificación de ADN es la prueba de paternidad y las pruebas de parentesco en el linaje materno o paterno con marcadores de ADN autosómicos, mitocondriales o del cromosoma Y. En muchos aspectos, los métodos de caracterización de ADN han aumentado la capacidad de resolver problemas de relaciones familiares y casos forenses que de otro modo podrían haber quedado sin resolver (El-Alfy *et al.*, 2012).

Los loci de ADN que se utilizan para genética forense (figura 1) deben de tener algunas propiedades claves que idealmente deberían ser según Goodwin *et al.* (2011) las siguientes:

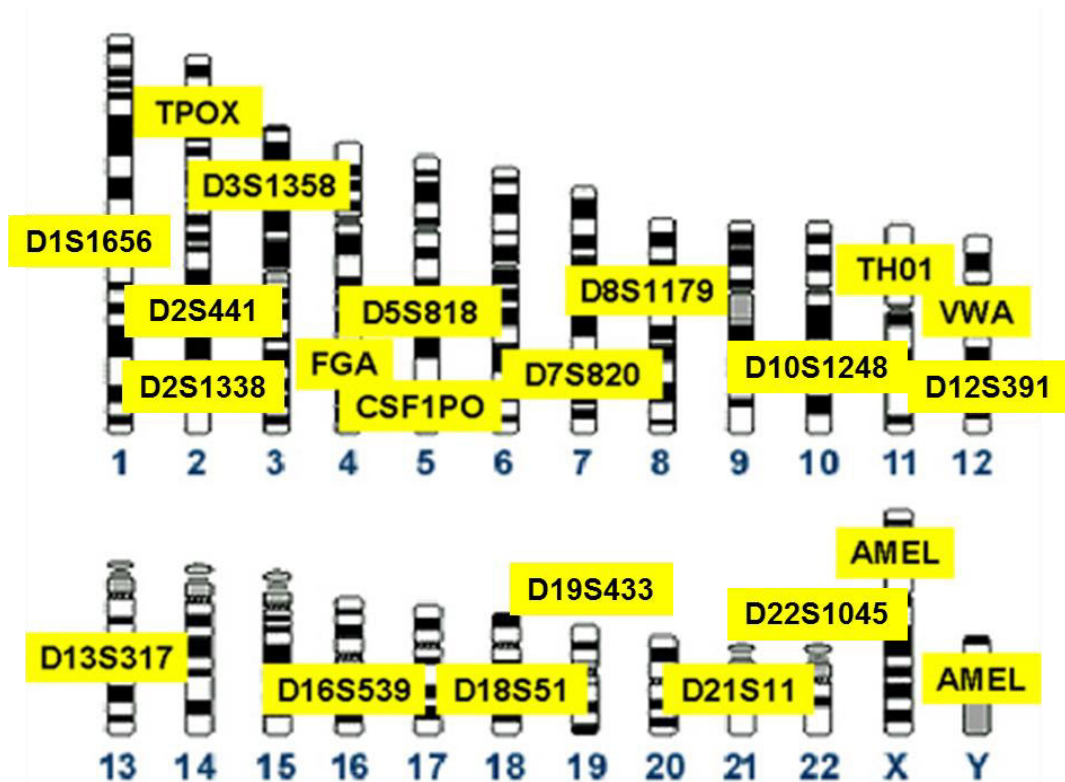
- Ser altamente polimórficos (variar ampliamente entre individuos).
- Ser económicos y sencillos de ser caracterizados.
- Proporcionar perfiles que sean simples de interpretar y fáciles de comparar entre laboratorios.
- No estar bajo ninguna presión selectiva.
- Tener una baja tasa de mutación.

Los STRs se usaron por primera vez en trabajos forenses a principios de la década

de 1990 (Goodwin *et al.*, 2011). En noviembre de 1997, en los Estados Unidos, el Buró Federal de Investigaciones (FBI) seleccionó 13 marcadores STR para servir como *coré* de su base de datos de ADN nacional denominado Sistema de Indexación del DNA Combinado (CODIS) (El-Alfy *et al.*, 2012). Estos marcadores fueron CSF1PO, FGA, TH01, TPOX, VWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51 y D21S11, y el locus amelogénico (AMEL- X/Y) (figura 2), el cual permite diferenciar el sexo (Ruitberg *et al.*, 2001; Butler *et al.*, 2004; Ziętkiewicz *et al.*, 2012). En el 2011 el FBI propuso implementar nuevos marcadores STRs en el programa CODIS, siendo los objetivos de esta expansión: el de reducir el número de coincidencias accidentales, acrecentar la compatibilidad internacional y aumentar el poder de discriminación para casos criminales y de personas desaparecidas (Hares, 2015). El 1 de Enero del 2017 el FBI agregó 7 nuevos marcadores STRs al CODIS (FBI, 2018) siendo estos D1S1656, D2S441, D2S1338, D10S1248, D12S391, D19S433, D22S1045 (figura 2).



**Figura 1.** Estructura de una Repetición Corta en Tándem (STR). Tomada de Hameed *et al.*, 2014.



**Figura 2.** Ubicación cromosómica de los 20 marcadores genéticos microsatélites del CODIS y el marcador genético del sexo, amelogénina. Figura tomada y editada de la dirección electrónica: <https://strbase.nist.gov/fbicore.htm>.

## 2.1. Características de los STRs

La comunidad forense ha seleccionado los loci STR basándose en ciertos criterios que cumplen las siguientes características (Gill *et al.*, 1996; Butler, 2005; Goodwin *et al.*, 2011):

- Alelos discretos y distinguibles
- La amplificación del locus debe de ser robusta y reproducible cuando se amplifican conjuntamente con otros marcadores.
- Alto poder de discriminación, usualmente  $>0.9$ , con una heterocigosidad observada  $>70\%$ .
- Ausencia de ligamiento genético con otros loci que se analizan.
- Separada localización cromosómica para asegurar que los loci estrechamente vinculados no sean elegidos.



- Bajo nivel de producción de *stutter* y bajos niveles de formación de artefactos durante la amplificación.
- Baja tasa de mutación.
- La longitud prevista de los alelos se debe de encontrar en el rango de 90-500 pb.

## **2.2. Clasificación de los STRs**

Los STRs son clasificados en varias categorías basándose como señala Butler (2005) en su “patrón de repetición”:

- Repeticiones simples: Contiene unidades de repetición que son idénticas en secuencia y longitud.
- Repeticiones compuestas: Comprenden dos o más repeticiones simples contiguas.
- Repeticiones complejas: Pueden contener varios bloques de repetición de variable unidad de longitud, como secuencias intermedias variables.
- Repeticiones Hipervariables complejas: Se distinguen por poseer una gran cantidad de alelos no consensuados que se diferencian en su secuencia así como en su tamaño, y por ende, son un reto para la reproducibilidad del genotipo. Esta categoría no se usa tan recurrentemente en la tipificación forense del ADN por que presenta inconvenientes en la nomenclatura de los alelos y la variabilidad de medición entre laboratorios.

## **2.3. Nomenclatura de los STRs**

La nomenclatura para los marcadores de ADN es bastante directa (Butler, 2005; Hameed *et al.*, 2014). Los marcadores de ADN que se encuentran fuera

de las regiones génicas se pueden designar por su posición cromosómica. Los loci STR D5S818 y DYS19 son ejemplos de marcadores que no se encuentran dentro de las regiones de genes. En estos casos, la "D" significa ADN (por su denominación en inglés "*Deoxyribonucleic acid*"). El siguiente carácter designa al cromosoma donde se encuentra el marcador, así el número "5" se refiere al cromosoma 5 e "Y" al cromosoma Y. La "S" se refiere al hecho de que el marcador de ADN es una secuencia de copia única. El número final indica el orden en que el marcador fue descubierto y categorizado para un cromosoma en particular. Así para el marcador de ADN D3S1358: D=DNA, 3=Cromosoma 3, S=Secuencia de copia única, 1358=Numero de locus descrito en el cromosoma 3. De otro lado, si un marcador es parte de un gen o se encuentra dentro de un gen, el nombre del gen se utiliza en la designación. Por ejemplo, el marcador STR TH01 es del gen de la tirosina hidroxilasa humana, ubicado en el cromosoma 11. La porción "01" de TH01 proviene del hecho de que la región de repetición en cuestión se encuentra dentro del intrón 1 del gen de la tirosina hidroxilasa. En este estudio se analizaron 20 marcadores que se describen a continuación (tabla 1) (Butler y Hill, 2012):

El STR D3S1358 (GenBank AC099539) es una repetición tetranucleotídica compuesta encontrada en el brazo corto del cromosoma 3 (3p21.31). Este locus posee tanto unidades repetitivas TCTA como TCTG con un rango alélico de 6 a 26 repeticiones (Butler y Hill, 2012).

El STR vWA (GenBank M25858) es una repetición tetranucleotídica compuesta encontrada en el intrón 40 del gen del Factor von Willebrand en el brazo corto del cromosoma 12 (12p13.31). Posee una repetición TCTA intercalada con una repetición TCTG, con un rango alélico de 10 a 25 repeticiones (Butler y Hill, 2012).

El STR D16S539 (GenBank AC024591) es una repetición tetranucleotídica simple encontrada en el brazo largo del cromosoma 16 (16q24.1). Este locus posee una unidad de repetición central de GATA con un rango de 4 a 17 unidades de longitud (Butler y Hill, 2012).

El STR CSF1PO (GenBank X14720) es una repetición simple tetranucleotídica encontrada en el sexto intrón del proto-oncogene c-fms para el receptor CSF-1 del brazo largo del cromosoma 5 (5q33.1). Contiene una repetición central AGAT y un rango en tamaño de 5 a 17 repeticiones (Butler y Hill, 2012).

El STR TPOX (GenBank M68651) es una repetición simple tetranucleotídica encontrada en el intrón 10 del gen de la peroxidasa tiroidea humana cerca del final del brazo corto del cromosoma 2 (2p25.3). Este locus posee una repetición simple AATG con un rango alélico que varían en tamaño de 4 a 16 repeticiones (Butler y Hill, 2012).

El STR D8S1179 (GenBank AF216671) es una repetición tetranucleotídica compuesta encontrada en el cromosoma 8 (8q24.13). El locus consta principalmente de alelos que contienen TCTA, aunque una unidad de repetición de TCTG ingresa al motivo para todos los alelos de más de 13 repeticiones, generalmente en la segunda o tercera posición desde el extremo 5' de la región de repetición. Tiene un rango alélico en tamaño de 6 a 20 repeticiones (Butler y Hill, 2012).

El STR D21S11 (GenBank AP000433) es una repetición tetranucleotídica compleja encontrada en el brazo largo del cromosoma 21 (21q21.1). Un número variable de bloques de repetición TCTA y TCTG rodea una sección

constante de 43 pb formada por la secuencia {[TCTA]<sub>3</sub> TA [TCTA]<sub>3</sub> TCA [TCTA]<sub>2</sub> TCCA TA}. Con un rango alélico que varía en tamaño de 12 a 43.2 repeticiones. Los alelos microvariantes x.2 surgen principalmente de una inserción de 2 pb (TA) en el extremo 3' de la región repetida (Butler y Hill, 2012).

El STR D18S51 (GenBank AP001534) es una repetición simple tetranucleotídica AGAA encontrada en el brazo largo del cromosoma 18 (18q21.33). Los alelos varían en tamaño desde 5.3 hasta 40 repeticiones. Existen varias variantes alélicas de x.2 debido a una delección de 2 pb de una pérdida de AG en la región del flanco 3' (Butler y Hill, 2012).

El STR D2S441 (GenBank AC079112) es una repetición tetranucleotídica localizada en el brazo corto del cromosoma 2(2p14) a más de 60 Mb de TPOX. Es un motivo de repetición compuesto con secuencias TCTA y TCAA pudiendo variar de 8 a 17 repeticiones (Butler y Hill, 2012).

El STR D19S433 (GenBank AC008507) es una repetición tetranucleotídica compuesta localizada en el cromosoma 19 (19q12) con alelos observados que van desde 5.2 hasta 20 repeticiones. Los alelos x.2 se deben a una eliminación AG antes de la repetición AAGG central (Butler y Hill, 2012).

El STR TH01 (GenBank D00269) es una repetición tetranucleotídica simple encontrada en el intrón 1 del gen tirosina hidroxilasa en el brazo corto del cromosoma 11 (11p15.5). TH01 tiene una secuencia de tetranucleótidos simple con un motivo de repetición de TCAT en la cadena superior en la secuencia de referencia de GenBank con alelos que varían en tamaño de 3 a 14 repeticiones (Butler y Hill, 2012).

El STR FGA (GenBank M64982) es una repetición tetranucleotídica compuesta encontrada en el tercer intrón del locus del fibrinógeno alfa humano en el brazo largo del cromosoma 4 (4q31.3). El locus contiene una repetición CTTT flanqueada en cada lado por repeticiones degeneradas. Los alelos reportados varían en tamaño desde 12.2 repeticiones hasta 51.2 repeticiones, abarcando más de 35 repeticiones. Una delección de 2 pb, de la pérdida de un CT, en la región justo antes del motivo de repetición del núcleo es responsable de los alelos microvariantes x.2 que son muy prevalentes en este sistema STR (Butler y Hill, 2012).

El STR D22S1045 (GenBank AL022314) es una repetición simple trinucleotídica, encontrada en el cromosoma 22 (22q12.3) con un rango alélico de 7 a 20 repeticiones ATT (Butler y Hill, 2012).

El STR D5S818 (GenBank AC008512) es una repetición simple tetranucleotídica encontrada en el brazo largo del cromosoma 5 (5q23.2). El locus posee unidades de repetición AGAT con alelos que varían en tamaño de 4 a 29 repeticiones (Butler y Hill, 2012).

El STR D13S317 (GenBank AL353628) es una repetición tetranucleotídica simple TATC que se encuentra en el brazo largo del cromosoma 13 (13q31.1). Los alelos comunes contienen entre 5 y 17 unidades de repetición (Butler y Hill, 2012).

El STR D7S820 (GenBank AC004848) es una repetición simple de tetranucleótidos que se encuentra en el brazo largo del cromosoma 7

(7q21.11). El locus posee principalmente una repetición GATA con alelos que varían en tamaño de 5 a 16 repeticiones (Butler y Hill, 2012).

El STR D10S1248 (GenBank AL391869) es una repetición simple de tetranucleótidos que se encuentra en el brazo largo del cromosoma 10 (10q26.3) y posee de 7 a 19 repeticiones de GGAA (Butler y Hill, 2012).

El STR D1S1656 (GenBank G07820) es una repetición tetranucleotídica que se encuentra en el brazo largo del cromosoma 1 (1q42) con alelos que van de 8 a 20.3 repeticiones. Los alelos x.3 surgen de una inserción de TGA típicamente después de cuatro repeticiones de TAGA completas (Butler y Hill, 2012).

El STR D12S391 (GenBank G08921) es un tetranucleótido compuesto que posee unidades de repetición AGAT/AGAC, es altamente polimórfico y se encuentra en el brazo corto del cromosoma 12 (12p13.2), a 6.3 megabases de VWA. Posee más de 50 alelos diferentes de 13 a 27.2 repeticiones de longitud. Se han identificado varios alelos de secuencia diferente pero del mismo tamaño a través del análisis de secuenciamiento (Butler y Hill, 2012).

El STR D2S1338 (GenBank AC010136) es una repetición tetranucleotídica compuesta que posee unidades de repetición TGCC/TTCC en el brazo largo del cromosoma 2 (2q35). Se han observado alelos que van de 10 a 31 repeticiones (Butler y Hill, 2012).

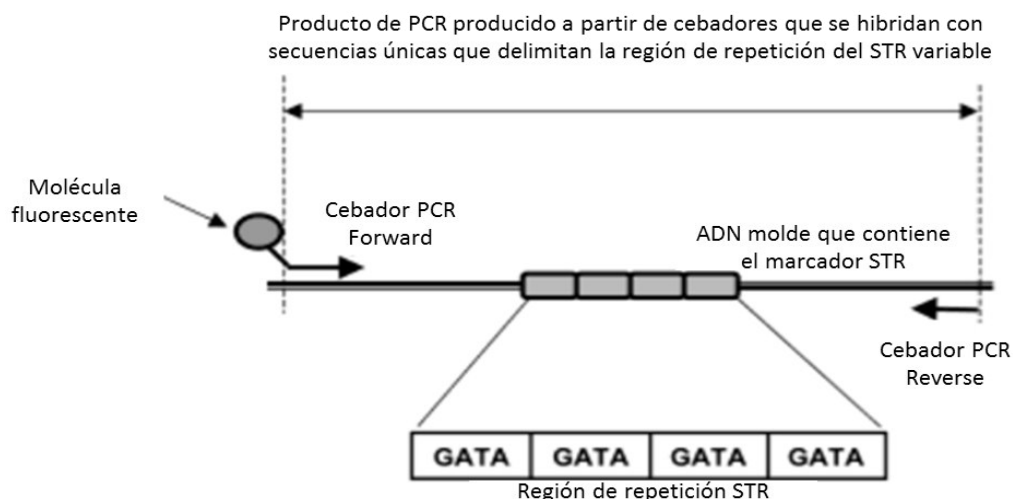
**Tabla 1.** Cuadro resumen de los 20 marcadores STRs autosómicos. Modificado de Butler y Hill, 2012.

<b>Locus</b>	<b>Localización Cromosómica</b>	<b>Posición Física</b>	<b>Acceso a GenBank</b>	<b>Categoría y motivo de Repetición</b>
<b>D3S1358</b>	3p21.31	Chr 3 45.582 Mb	AC099539	compuesta TCTA/TCTG
<b>vWA</b> intrón 40th, Factor von Willebrand	12p13.31	Chr 12 6.093 Mb	M25858	compuesta TCTA/TCTG
<b>D16S539</b>	16q24.1	Chr. 16 86.386 Mb	AC024591	simple GATA
<b>CSF1PO</b> intrón 6 th, protooncogene c-fms	5q33.1	Chr 5 149.455 Mb	X14720	simple AGAT
<b>TPOX</b> intrón 10th, peroxidasa tiroidea	2p25.3	Chr 2 1.493 Mb	M68651	simple AATG
<b>D8S1179</b>	8q24.13	Chr 8 125.907 Mb	AF216671	compuesta TCTA/TCTG
<b>D21S11</b>	21q21.1	Chr 21 20.554 Mb	AP000433	compleja TCTA/ TCTG
<b>D18S51</b>	18q21.33	Chr 18 60.949 Mb	AP001534	simple AGAA
<b>D2S441</b>	2p14	Chr 2 68.239 Mb	AC079112	compuesta TCTA/TCAA
<b>D19S433</b>	19q12	Chr 19 30.416 Mb	AC008507	compuesta AAGG/TAGG
<b>TH01</b> intrón 1st, tirosina hidroxilasa	11p15.5	Chr 11 2.192 Mb	D00269	simple TCAT
<b>FGA</b> intrón 3rd, fibrinógeno alfa	4q31.3	Chr 4 155.509 Mb	M64982	compuesta CTTT/TTCC
<b>D22S1045</b>	22q12.3	Chr 22 37.536 Mb	AL022314	simple ATT
<b>D5S818</b>	5q23.2	Chr 5 123.111 Mb	AC008512	simple AGAT
<b>D13S317</b>	13q31.1	Chr 13 82.692 Mb	AL353628	simple TATC
<b>D7S820</b>	7q21.11	Chr 7 83.789 Mb	AC004848	simple GATA
<b>D10S1248</b>	10q26.3	Chr 10 131.093 Mb	AL391869	simple GGAA
<b>D1S1656</b>	1q42	Chr 1 230.905 Mb	G07820	compuesta TAGA
<b>D12S391</b>	12p13.2	Chr 12 12.450 Mb	G08921	compuesta AGAT/AGAC
<b>D2S1338</b>	2q35	Chr 2 218.879 Mb	AC010136	compuesta TGCC/TTCC

## 2.4. Electroforesis capilar

La Electroforesis Capilar (EC) es un tipo de separación de fragmentos por su tamaño al ser sometidos a electricidad, que consiste primero en amplificar los STRs con un cebador (*primer*) que tiene unido covalentemente en el extremo 5' una molécula fluorescente (figura 3), pudiendo ser detectada en tiempo real durante la corrida. Esta técnica usa un tubo de vidrio angosto (capilar) lleno de una solución polimérica enredada que permite separar las moléculas de ADN por tamaño y color de fluorescencia (Butler *et al.*, 2004; Goodwin *et al.*, 2011). Antes de la electroforesis, a los amplicones STRs se les debe agregar formamida desionizada y el estándar de tamaño del carril interno o estándar interno. La formamida desionizada permite desnaturalizar el ADN, para luego asegurar esta desnaturalización calentando las muestras a 95 °C (Goodwin *et al.*, 2011). De otro lado, el estándar interno que contiene fragmentos de ADN de tamaño conocido y marcados con un color de molécula fluorescente diferente, suele correrse junto con cada muestra para calibrar los tamaños de corrida a corrida (Butler *et al.*, 2004). Las muestras se transfieren al capilar mediante inyección electrocinética, se aplica un voltaje y los fragmentos de ADN amplificados y los estándares del tamaño del carril interno, migran al capilar. Después de la inyección, un constante voltaje es aplicado a través del capilar y los productos de PCR migran hacia el ánodo cargado positivamente, viajando a través del polímero que llena el capilar y actúa como matriz de tamizado (Goodwin *et al.*, 2011).

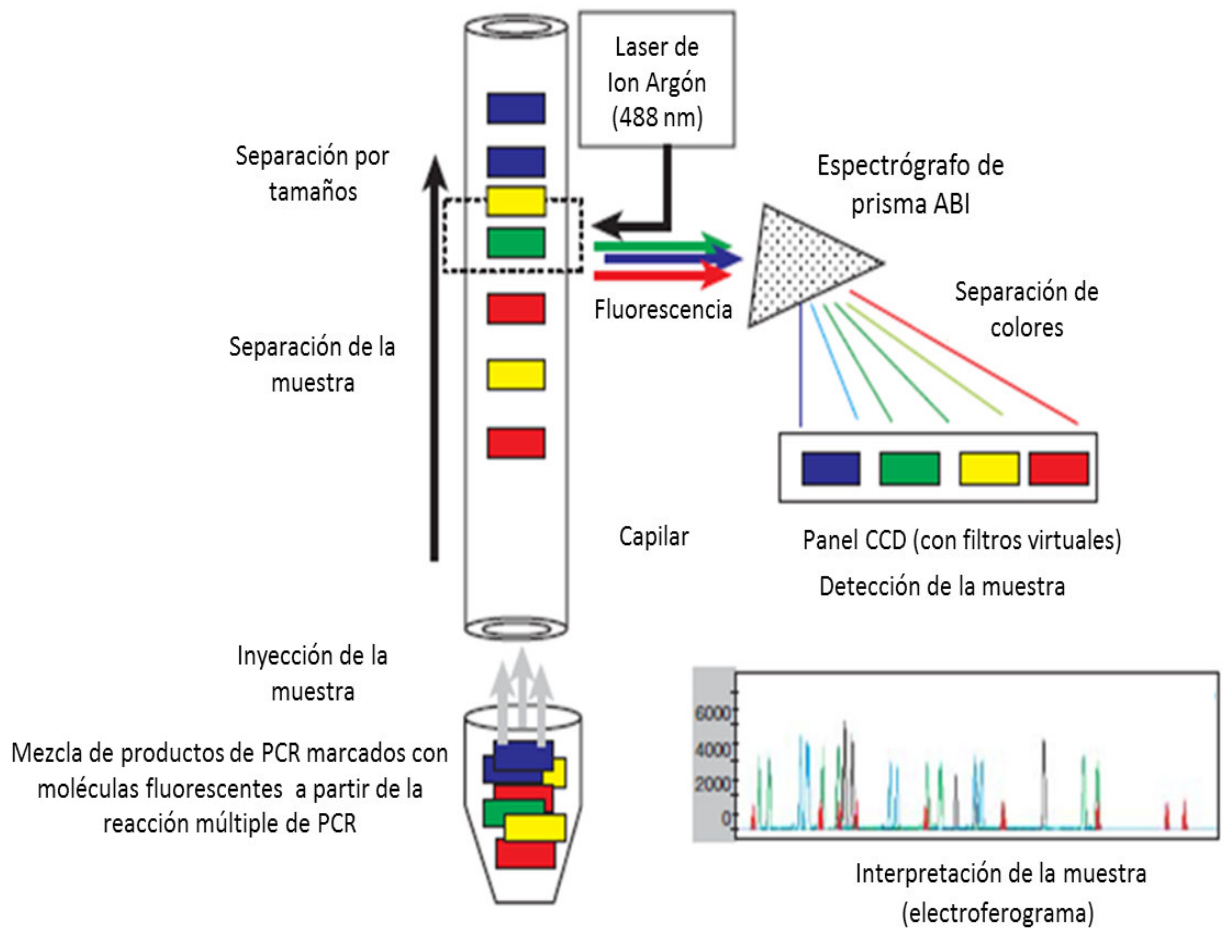




**Figura 3.** Esquema de las posiciones del cebador de PCR para la amplificación de un marcador STR de ADN. Las flechas de una sola cabeza representan las posiciones del cebador. La flecha de dos cabezas ilustra el tamaño total del producto de PCR usando un conjunto particular de cebadores. El tamaño del producto de PCR se mide y se convierte de nuevo en el número de unidades repetidas presentes en la muestra para fines de genotipado. Una molécula fluorescente está presente en uno de los cebadores para marcar el producto de PCR con un color específico. Butler *et al.*, 2004.

Durante el período de electroforesis, un láser de ion argón se ilumina a través de una pequeña ventana de vidrio en el capilar (figura 4), y cuando los productos de PCR marcados con las moléculas fluorescentes viajan más allá de la ventana, son excitados por el láser y emiten fluorescencia, que se detecta mediante una cámara CCD (*charged coupled device*), y luego son grabados por el software de recolección (Goodwin *et al.*, 2011). Estos datos son recogidos en forma de electroferogramas multicolores y el software determina automáticamente los tamaños de los alelos STR basándose en una curva estándar producida a partir del estándar de tamaño interno. El genotipado de STR se realiza comparando los tamaños de alelos en cada muestra con los tamaños de los alelos presentes en una escalera alélica, que contiene alelos comunes que se han secuenciado previamente (Butler *et al.*, 2004). Así la electroforesis de una muestra tarda hasta 30 minutos, después de lo cual el polímero en el capilar se reemplaza con polímero fresco y la siguiente muestra puede analizarse (Goodwin *et al.*, 2011).

La EC ofrece una serie de ventajas sobre las separaciones de gel convencional en términos de velocidad, resolución, sensibilidad y manejo de datos (Smith y Nelson, 2004). En primer lugar, los pasos de inyección, separación y detección se pueden automatizar por completo, permitiendo que múltiples muestras se ejecuten sin supervisión (Butler, 2005). Además, solo se consumen cantidades mínimas de muestra en el proceso de inyección, en el rango de 1 nanogramo (Butler, 2005; Smith y Nelson, 2004) y las muestras se pueden volver a analizar fácilmente si es necesario. Esta es una ventaja importante para especímenes forenses apreciados que a menudo no pueden ser reemplazados fácilmente. La separación en los capilares se puede realizar en minutos en lugar de horas debido a los voltajes más altos que se permiten con la disipación de calor mejorada de los capilares. Otra ventaja es que la información cuantitativa está fácilmente disponible en un formato electrónico después de la finalización de una corrida, por lo que no se requieren pasos adicionales, como escanear el gel o tomar una fotografía (Butler, 2005).



**Figura 4.** Ilustración esquemática de la separación y detección de alelos STRs con un analizador genético ABI Prism 310. Modificado del libro: *Forensic DNA typing: biology, technology, and genetics of STR markers*. (Butler, 2005).

### **3. HIPÓTESIS Y OBJETIVOS**

#### **3.1. Hipótesis**

Existen diferencias significativas entre las frecuencias alélicas de 20 marcadores STRs, utilizadas para la identificación de individuos, en una muestra poblacional mestiza peruana y las registradas en la población Hispanoamericana.

##### **3.1.1. Hipótesis Estadísticas**

**Ho:** Existen diferencias significativas entre las frecuencias alélicas de 20 loci STRs reportados para la población hispanoamericana y las frecuencias alélicas de la muestra poblacional en estudio.

**Ha:** No existen diferencias significativas entre las frecuencias alélicas de 20 loci STRs reportados para la población hispanoamericana y las frecuencias alélicas de la muestra poblacional en estudio.

#### **3.2. Objetivo General**

Determinar si las frecuencias alélicas de 20 marcadores STRs para la identificación humana en una muestra poblacional mestiza del Perú son estadísticamente diferentes al de la población hispanoamericana.

##### **3.2.1. Objetivos Específicos**

- Determinar las frecuencias alélicas de 20 marcadores STRs para la identificación humana en una muestra poblacional mestiza del Perú.

- Establecer las variaciones alélicas de 20 loci STRs en una muestra poblacional mestiza peruana.
- Analizar las frecuencias alélicas de 20 loci STRs en una muestra poblacional mestiza peruana.
- Verificar si la muestra poblacional en estudio se encuentra en equilibrio genético de Hardy-Weinberg.
- Comparar las frecuencias alélicas de 20 loci STRs en una muestra poblacional mestiza del Perú con las reportadas para la población hispanoamericana.

## **4. MATERIALES Y MÉTODOS**

### **4.1. Población de Estudio**

La población analizada en este estudio estuvo conformada por una muestra poblacional mestiza del Perú constituida por 200 personas voluntarias no emparentadas entre sí de ambos sexos (103 hombres y 97 mujeres), reclutados en la provincia de Lima.

A cada participante se le brindó información detallada acerca de los objetivos del estudio. Los voluntarios que accedieron a participar firmaron el Consentimiento Informado (Anexo 1) aprobado por el Comité Institucional de Ética en Investigación del Instituto de Medicina Tropical “Daniel Alcides Carrión” de la Universidad Nacional Mayor de San Marcos (Anexo 2, Constancia de Aprobación CIEI-2018-015).

Todos los participantes cumplían los siguientes criterios.

### **Criterios de inclusión**

- Voluntarios que hayan nacido y/o habiten en Lima o Callao.
- Mayores de 18 años.
- Que acepten participar en el estudio a través del consentimiento informado.

### **Criterios de Exclusión**

- Voluntarios que no hayan nacido y/o habiten en Lima o Callao.
- Que hayan recibido transfusión sanguínea.
- Haber recibido trasplante de medula ósea.
- No mayores de 18 años.

## **4.2. Materiales**

### **4.2.1. Reactivos**

- Kit de amplificación de PCR VeriFiler Express.
- Polímero POP-4™ 3500/3500xL.
- Formamida Hi-Di™.
- GeneScan 600 LIZ Size Standard v2.0.
- Buffer TE.

### **4.2.2. Consumibles**

- Tarjetas COPAN NUCLEIC- CARD™.
- Lancetas.
- Puntas de micropipeta sin filtro de 10, 200 y 1000 microlitros.

- Tubos de centrifuga de 0.2, 1.5 microlitros.
- Placa óptica de 96 pocillos MicroAmp™.
- Film adhesivo transparente MicroAmp™.
- Matriz Capilar de Analizador Genético 3500xL.
- Septa 3500/3500XL de 96 pocillos.

#### **4.2.3. Equipos**

- Micropipetas Eppendorf® graduables hasta 10, 200 y 1000 microlitros.
- Micropipeta multicanal.
- Puncher.
- Centrifuga.
- Termociclador Veriti™ 96 pocillos.
- Analizador Genético ABI 3500 XL.

### **4.3. Métodos**

#### **4.3.1. Toma de muestra**

La toma de muestra de sangre se realizó mediante la punción en la yema del dedo con lancetas estériles, colocando la muestra de sangre mediante goteos en la tarjeta COPAN NUCLEIC- CARD™; se dejó secar a temperatura ambiente. Luego se colocó cada tarjeta en un sobre, previamente esterilizado y codificado según las pautas establecidas. Los sobres se guardaron a temperatura ambiente en un recipiente cerrado hasta su posterior utilización.

#### 4.3.2. Amplificación de ADN para Electroforesis Capilar

La reacción en cadena de la polimerasa (PCR) se llevó a cabo en el Termociclador Veriti™ de 96 pocillos, siguiendo las instrucciones de manufactura (VeriFiler™ Express PCR Amplification Kit User Guide. 2017), mediante el siguiente resumen.

- a) Preparación del mix por muestra para la amplificación con el kit VeriFiler Express: Se agregó 5  $\mu$ l de Master mix, 5  $\mu$ l de Primer Set y 2.5  $\mu$ l de Buffer TE<sup>-4</sup>.

Reactivos	Volumen por Reacción ( $\mu$ L)
Master Mix	5
Primer Set	5
Buffer TE <sup>-4</sup>	2.5

- b) Se depositó 12.5  $\mu$ l del Mix a cada pocillo de la placa de reacción de 96 pocillos MicroAmp™.
- c) Previamente las tarjetas COPAN NUCLEIC- CARD™ se depositaron en la incubadora a 60 °C por una hora, realizándose luego la perforación de ellas con el micropuncher, obteniendo discos de 1.2 mm de diámetro que fueron depositados en la placa de reacción de 96 pocillos Microamp™. Para el control negativo se agregó un disco perforado de 1.2 mm de diámetro en blanco (sin muestra). Para el control positivo se agregó 3  $\mu$ l de control ADN 007 (reactivo perteneciente al kit).
- d) Se selló la placa con el film adhesivo transparente MicroAmp.
- e) Se centrifugo a 3000 rpm por 20 segundos.
- f) Se realizó la amplificación en el Termociclador Veriti™ 96 pocillos, según el protocolo del kit.



Paso inicial de incubación	25 ciclos			Extensión Final	Final
	Desnaturalizar	Annealing	Extensión		
Hold	Ciclo			Hold	Hold
95 °C 1 minuto	94 °C 20 segundos	59 °C 16 segundos	65 °C 29 segundos	60 °C 5 minutos	4 °C .....

#### 4.3.3. Electroforesis Capilar

Se procedió a realizar la electroforesis capilar en el analizador genético ABI 3500 XL de Applied Biosystems. Se siguió las instrucciones de manufactura.

- a) Preparación del mix por muestra: 9.6  $\mu$ L de formamida Hi-Di™ 0.4  $\mu$ L de GeneScan™ 600 LIZ®.

Reactivo	Volumen por Reacción ( $\mu$ L)
Formamida Hi-Di™	9.6
GeneScan™ 600 LIZ®	0.4

- b) Se depositó 10  $\mu$ L del Mix a cada pocillo de la placa de reacción de 96 pocillos MicroAmp™.
- c) Se agregó 1  $\mu$ L de cada muestra amplificada a cada pocillo de la placa de reacción de 96 pocillos MicroAmp™ y por cada 24 muestras 1  $\mu$ L de allelic ladder
- d) Se selló la placa con el septa 3500/3500XL de 96 pocillos, luego se realizó un breve homogenizado vortex y centrifugado a la placa.
- e) Se calentó la placa de reacción a 95 °C por 5 minutos.
- f) Inmediatamente se enfrió la placa en hielo por 3 minutos.
- g) Se colocó la placa en el analizador genético ABI 3500 XL, y se procedió a iniciar la corrida electroforética.

#### **4.3.4. Identificación de los alelos**

El análisis e identificación de los diferentes alelos para los distintos marcadores evaluados se realizó mediante el software Informático GeneMapper® ID Software v4.0 de Applied Biosystems.

#### **4.3.5. Métodos de Análisis e Interpretación de Resultados**

Con los datos obtenidos se realizaron los siguientes cálculos para cada uno de los marcadores estudiados.

##### **4.3.5.1. Parámetros de interés poblacional**

Los parámetros de interés poblacional fueron calculados utilizando el programa computacional Arlequín 3.5.2.2. (Excoffier y Lischer, 2010).

##### **4.3.5.1.1. Estimación de las frecuencias alélicas**

La frecuencia alélica es el cálculo de la proporción relativa de alelos de una determinada población. Dichas frecuencias se miden en tantos por uno o en tantos por ciento, calculándose por medio del recuento de cada alelo y dividiéndose por el total de alelos estudiados. Así, la sumatoria de las frecuencias alélicas siempre será 1 (o 100%) (Filho, 2012).

##### **4.3.5.1.2. Índice de Heterocigosidad**

Este índice representa la probabilidad de que en una población se tome al azar dos alelos del mismo locus y estos sean distintos (García Fernández, 1999 citado en Barletta, 2010). La heterocigosidad observada ( $H_{obs}$ ) manifiesta la frecuencia de individuos heterocigotos

hallados en la población de estudio, por otro lado la heterocigosidad esperada ( $H_{esp}$ ) o diversidad génica, es una valoración más exacta de la heterocigosidad en la población (Filho, 2012), ya que según Nei y Roychoudhury (1974) esta representa la proporción de heterocigotos que se desearía hallar en la población si esta estuviera en equilibrio de Hardy-Weinberg.

Se determina mediante la siguiente fórmula:

$$He = \frac{n}{n-1} \left( 1 - \sum p_i^2 \right)$$

Siendo  $n$  el número de cromosomas muestreados,  $he$  la heterocigosidad esperada, y  $P_i$  la frecuencia del alelo o haplotipo “i” (Nei, 1987).

#### **4.3.5.1.3. Equilibrio de Hardy- Weinberg**

Dicho equilibrio de Hardy-Weinberg (H-W) predice la estabilidad de las frecuencias alélicas y genotípicas de una generación a la siguiente. El primer objetivo en esta prueba es determinar si alelos dentro de un locus son independientes el uno con el otro. Se considera que un locus está en equilibrio de H-W cuando las frecuencias genotípicas y alélicas en una población permanecen constantes a través del tiempo; sin embargo, poblaciones naturales usualmente violan el equilibrio de H-W en algunos niveles y por lo tanto causan que las frecuencias alélicas cambien a través del tiempo (Butler, 2005).

El equilibrio de H-W asume los principios de un tamaño poblacional infinito en el cual los apareamientos ocurren de manera aleatoria, además incluye la no presencia de selección natural y deriva génica, con ausencia de migración o mutación que pueda introducir nuevos alelos (Butler, 2005).

Otro objetivo de la prueba del equilibrio de H-W es la búsqueda de alguna indicación de exceso de homocigotos. La principal explicación para este suceso es por la presencia de alelos nulos, lo cual es debido a una mutación en el sitio de unión del cebador, en donde únicamente un alelo es amplificado y observado en un individuo heterocigoto verdadero (Butler, 2005).

Para determinar la desviación del equilibrio de H-W se utiliza la prueba estadística del chi-cuadrado ( $X^2$ ) (Goodwin *et al.*, 2011), el cual es establecido sumando el valor al cuadrado de la diferencia entre los resultados observados (obs) y los resultados esperados (esp) dividido por los resultados esperados (Butler, 2005).

$$X^2 = \sum_{i=1}^k \frac{(obs_i - esp_i)^2}{esp_i}$$

El valor de chi cuadrado resultante se compara con una tabla de números para ver si hay una desviación significativa de los valores "normales" esperados. Los

valores altos de chi cuadrado indican discrepancias entre los resultados observados y esperados. Se pueden aplicar diferentes "grados de libertad" a los datos según la situación (Butler, 2005).

#### **4.3.5.1.4. Frecuencias Mínimas**

Existen algunos alelos que no son muestreados suficientemente y que una estimación de su frecuencia alélica es incierta en el sentido de considerar que el alelo es tan raro que se representa solamente una o varias veces en un conjunto de datos (Hameed *et al.*, 2015). Por esta razón Chakraborty (1992), señaló el uso de una frecuencia de alelo mínima para asegurar que un alelo ha sido muestreado lo suficiente como para ser utilizado confiablemente en pruebas estadísticas. Además, el informe del *National Research Council* (1996) establece que una estimación de la frecuencia de un alelo puede ser muy imprecisa si el alelo es tan raro que solo se representa una o varias veces en una base de datos; y algunos alelos raros podrían no estar representados en absoluto (Butler, 2005). Un alelo raro es definido como cualquier alelo que posee una frecuencia inferior a 0,01 en una base de datos de muestra poblacional (Restrepo *et al.*, 2011). Butler (2005) recomienda que cada alelo se observe al menos cinco veces para incluirlo en cálculos estadísticos confiables. La frecuencia mínima de alelos es por lo tanto  $5/2N$ , donde  $N$  es el número de individuos muestreados de una población y  $2N$  es el número de

cromosomas contados porque los autosomas están en pares debido a la herencia de un alelo de la madre y uno del padre.

#### 4.3.5.1.5. Desequilibrio de Ligamiento

La capacidad de combinar información de múltiples loci refuerza las estadísticas de las probabilidades de coincidencias y conduce a un poder real para la tipificación de STRs (Butler, 2005). Para que las probabilidades de coincidencia aleatoria se combinen desde múltiples loci aplicando la regla del producto (Martínez-Sevilla *et al.*, 2016), se requiere que esos loci sean independientes entre sí, es decir que estas regiones de ADN se encuentran en equilibrio de ligamiento (Butler, 2005), por otro lado el desequilibrio de ligamiento se refiere a las dependencias entre alelos en diferentes loci, y se puede definir como la asociación no aleatoria de alelos en haplotipos (Tillmar, 2010).

El clásico coeficiente de desequilibrio de ligamiento que mide la desviación de la asociación aleatoria entre alelos de diferentes loci es expresado como (Lewontin y Kojima, 1960):

$$D_{ij} = p_{ij} - p_i p_j$$

Donde  $p_{ij}$  es la frecuencia del haplotipo que tiene el alelo  $i$  en el primer locus y el alelo  $j$  en el segundo locus, y  $p_i$  y  $p_j$  son las frecuencias de los alelos  $i$  y  $j$ , respectivamente.

#### 4.3.5.2. Parámetros de Interés forense

Los parámetros de interés forense fueron calculados utilizando el programa PowerStats V12 (Tereba, 1999).

##### 4.3.5.2.1. Índice de Contenido Polimórfico (PIC)

Este parámetro indica el nivel polimórfico de un locus. El PIC también se calcula usando frecuencias alélicas de marcadores a través de la siguiente ecuación (Botstein *et al.*, 1980):

$$PIC = 1 - \sum_{i=1}^n p_i^2 - \sum_{i=1}^{n-1} \sum_{j=i+1}^n 2p_i^2 p_j^2$$

Donde  $n$  es el número de alelos y  $p_i$  es la probabilidad del alelo  $i^{th}$ .

##### 4.3.5.2.2. Probabilidad de Coincidencia (PC)

Se define como la posibilidad de que dos individuos tomados al azar de la misma población concuerden en su genotipo para ese locus. Esta probabilidad fue descrita por Jones (1972) y se calcula a partir de la siguiente fórmula:

$$PM = \sum_{i=1}^n p_i^2$$

Siendo  $p_i$  la frecuencia esperada de cada genotipo.

La probabilidad de coincidencia combinada para más de un locus es el producto de la probabilidad de coincidencia individual para cada locus, asumiendo que ellos no están ligados (Huston, 1998).

#### 4.3.5.2.3. Poder de Discriminación (PD)

Es la probabilidad de que dos individuos seleccionados al azar y no relacionados, logren distinguirse genéticamente por medio del estudio de un marcador o un conjunto de marcadores. El poder de discriminación depende del polimorfismo de cada locus analizado, así como del número de locus estudiados (Filho, 2012):

$$PD = 1 - PC$$

Donde  $PC$  es probabilidad de coincidencia (Fisher, 1951).

El poder de discriminación combinado para múltiples loci es calculado usando la siguiente fórmula (Huston, 1998).

$$PD_{combinado} = 1 - \prod_{i=1}^n (1 - PD_i)$$

#### 4.3.5.2.4. Poder de Exclusión (PE)

Es la probabilidad de descartar a un individuo que esté implicado en cierto asunto forense. Por ejemplo, cuando se trata de asuntos de paternidad, vendría a ser la probabilidad de descartar a un “falso padre”. Así el PE constituye la proporción de sujetos implicados falsamente en un peritaje (Filho, 2012).

Ohno *et al.* (1982) establecieron la siguiente fórmula para calcular el PE en estructuras autosómicas:



$$CE = \sum_{i=1}^n p_i (1 - p_i)^2 (1 - p_i + p_i^2) + \sum_{i=1}^{n-1} \sum_{j=i+1}^n p_i p_j (p_i + p_j) (1 - p_i - p_j)^2$$

Siendo  $p_i$  y  $p_j$  frecuencias génicas respectivas.

La efectividad en la exclusión se logra utilizando un conjunto de marcadores genéticos la cual se denomina probabilidad acumulativa de exclusión (PEcum) o poder de exclusión combinado, el cual se calcula de la siguiente manera (Chakraborty y Schull, 1976):

$$PE_{cum} = 1 - (1 - P_1)(1 - P_2)(1 - P_3) \dots (1 - P_n)$$

Donde  $P_1, P_2, P_3, \dots, P_n$  representan las probabilidades de exclusión de cada locus analizado.

#### 4.3.5.3. Diferenciación poblacional

La estimación de la diferenciación poblacional se realizó mediante el uso del estadístico  $F_{st}$  y la distancia genética de Reynolds, por medio del empleo de las frecuencias alélicas de la muestra poblacional mestiza peruana y poblaciones mundiales. Se aplicó el programa PHYLIP 3.695 (Felsenstein, 2013). Dichas distancias genéticas obtenidas se representaron en un árbol UPGMA usando el programa MEGA 6.06 (Tamura *et al.*, 2013) y adicionalmente se graficó un HEATMAP con Microsoft Excel 2010.

#### 4.3.5.3.1. Índice de Fijación

Uno de los estadísticos más utilizados para caracterizar la diferenciación es  $F_{st}$ , que es uno de los tres índices de fijación introducidos por Sewall Wright (Gattepaille, 2015). El efecto de la subdivisión de la población se mide con el parámetro denominado “índice de fijación” (simbolizado por  $F_{st}$ ) (Sánchez, 2017), el cual según García (2008) “mide la reducción media de heterocigosis de una subpoblación relativa a la población total debido a la deriva genética entre subpoblaciones”. De esta manera, mide el nivel de diferenciación genética que existe entre “subpoblaciones”. El valor oscila entre 0 (no existe diferenciación) y 1 (diferenciación completa) (García, 2008).

$$F_{st} = \frac{H_T - H_s}{H_t}$$

En donde Piñero *et al.* (2008) detallan que “ $H_t$  es el promedio de la heterocigosis esperada en la población total para todos los locis y  $H_s$  es el promedio de la heterocigosis esperada dentro de subpoblaciones para todos los locis”.

## **5. RESULTADOS**

### **5.1. Marcadores STRs analizados en la muestra poblacional mestiza peruana**

#### **5.1.1. Marcador D3S1358**

Se observaron 9 alelos para este marcador (tabla 2), siendo los alelos con mayor frecuencia el 16 (0.2925) y 15 (0.4950), mientras que los alelos de menor frecuencia fueron el 11 (0.0025) y 12 (0.0025).

#### **5.1.2. Marcador vWA**

Se observaron 8 alelos para este marcador (tabla 2), siendo los alelos con mayor frecuencia el 16 (0.3450) y 17 (0.3775), mientras que los alelos de menor frecuencia fueron el alelo 13 (0.0025) y 20 (0.0100).

#### **5.1.3. Marcador D16S539**

Se observaron 8 alelos para este marcador (tabla 2), siendo los alelos con mayor frecuencia el 12 (0.2400) y el 10 (0.2575), mientras que los alelos de menor frecuencia fueron el alelo 8 (0.0025) y 15 (0.0025).

#### **5.1.4. Marcador CSF1PO**

Se observaron 10 alelos para este marcador (tabla 2), siendo los alelos con mayor frecuencia el alelo 11 (0.2650) y 12 (0.4025), mientras que los alelos de menor frecuencia fueron el alelo 7 (0.0025), 12.3 (0.0025), 15 (0.0025).

#### **5.1.5. Marcador TPOX**

Se observaron 8 alelos para este marcador (tabla 2), siendo los alelos con mayor frecuencia el alelo 11 (0.2350) y 8 (0.5850), mientras que los alelos

de menor frecuencia fueron el alelo 8.2 (0.0025), el alelo 6 (0.0050) y 13 (0.0050).

#### **5.1.6. Marcador D8S1179**

Se observaron 9 alelos para este marcador (tabla 2), siendo los alelos con mayor frecuencia el alelo 14 (0.2100) y 13 (0.3400), mientras que los alelos de menor frecuencia fueron el alelo 9 (0.0075) y 17 (0.0075).

#### **5.1.7. Marcador D21S11**

Se observaron 13 alelos para este marcador (tabla 2), siendo los alelos con mayor frecuencia el alelo 30 (0.2050) y 32.2 (0.2300), mientras que los alelos de menor frecuencia fueron el alelo 29.2 (0.0025), 32 (0.0025), 35 (0.0025).

#### **5.1.8. Marcador D18S51**

Se observaron 14 alelos para este marcador (tabla 2), siendo los alelos con mayor frecuencia el alelo 13 (0.1525) y 14 (0.2325), mientras que los alelos de menor frecuencia fueron el alelo 21 (0.0025) y 25 (0.0025).

#### **5.1.9. Marcador D2S441**

Se observaron 9 alelos para este marcador (tabla 2), siendo los alelos con mayor frecuencia el alelo 11 (0.2025) y 10 (0.5875), mientras que los alelos de menor frecuencia fueron el alelo 9.1 (0.0025) y 16 (0.0025).

#### **5.1.10. Marcador D19S433**

Se observaron 14 alelos para este marcador (tabla 2), siendo los alelos con mayor frecuencia el alelo 15 (0.1725) y 14 (0.3300), mientras que los alelos de menor frecuencia fueron el alelo 11.2 (0.0025) y 14.1 (0.0025).

#### **5.1.11. Marcador TH01**

Se observaron 5 alelos para este marcador (tabla 2), siendo los alelos con mayor frecuencia el alelo 6 (0.2500) y 7 (0.4500), mientras que los alelos de menor frecuencia fueron el alelo 8 (0.0425) y el 9 (0.0625).

#### **5.1.12. Marcador FGA**

Se observaron 12 alelos para este marcador (tabla 2), siendo los alelos con mayor frecuencia el alelo 24 (0.1825) y 25 (0.1825), mientras que los alelos de menor frecuencia fueron el alelo 22.2 (0.0025) y 28 (0.0025).

#### **5.1.13. Marcador D22S1045**

Se observaron 8 alelos para este marcador (tabla 2), siendo los alelos con mayor frecuencia el alelo 16 (0.4125) y 15 (0.4650), mientras que los alelos de menor frecuencia fueron el alelo 12 (0.0025) y 19 (0.0025).

#### **5.1.14. Marcador D5S818**

Se observaron 8 alelos para este marcador (tabla 2), siendo los alelos con mayor frecuencia el alelo 12 (0.1825) y 11 (0.5025), mientras que los alelos de menor frecuencia fueron el alelo 8 (0.0050) y 14 (0.0050).

#### **5.1.15. Marcador D13S317**

Se observaron 10 alelos para este marcador (tabla 2), siendo los alelos con mayor frecuencia el alelo 12 (0.1800) y 9 (0.3675), mientras que los alelos de menor frecuencia fueron el alelo 9.2 (0.0025), 15 (0.0025), 17 (0.0025).

#### **5.1.16. Marcador D7S820**

Se observaron 8 alelos para este marcador (tabla 2), siendo los alelos con mayor frecuencia el alelo 10 (0.2550) y 11 (0.3675), mientras que los alelos de menor frecuencia fueron el alelo 7 (0.0025) y 14 (0.0050).

#### **5.1.17. Marcador D10S1248**

Se observaron 9 alelos para este marcador (tabla 2), siendo los alelos con mayor frecuencia el alelo 15 (0.2500) y 14 (0.4275), mientras que los alelos de menor frecuencia fueron el alelo 11 (0.0025), 18 (0.0025).

#### **5.1.18. Marcador D1S1656**

Se observaron 14 alelos para este marcador (tabla 2), siendo los alelos con mayor frecuencia el alelo 17.3 (0.1850) y 16 (0.2175), mientras que los alelos de menor frecuencia fueron el alelo 18 (0.0025), 10 (0.0050), 19.3 (0.0050).

#### **5.1.19. Marcador D12S391**

Se observaron 17 alelos para este marcador (tabla 2), siendo los alelos con mayor frecuencia el alelo 20 (0.2650) y 19 (0.2725), mientras que los alelos de menor frecuencia fueron el alelo 14 (0.0025) y 19.1 (0.0025).

#### **5.1.20. Marcador D2S1338**

Se observaron 11 alelos para este marcador (tabla 2), siendo los alelos con mayor frecuencia el alelo 17 (0.1800), 23 (0.1800) y 19 (0.2250), mientras que los alelos de menor frecuencia fueron el alelo 26 (0.0075) y 21 (0.0225).

**TABLA 2.** Frecuencias alélicas y parámetros de interés poblacional.

ALELO	Marcadores																			
	D3S1358	vWA	D16S539	CSF1PO	TPOX	D8S1179	D21S11	D18S51	D2S441	D19S433	TH01	FGA	D22S1045	D5S818	D13S317	D7S820	D10S1248	D1S1656	D12S391	D2S1338
6					0.0050						0.2500									
7				0.0025							0.4500			0.1100		0.0025				
8			0.0025	0.0100	0.5850						0.0425			0.0050	0.0450	0.0425				
8.2					0.0025															
9			0.2075	0.0075	0.0300	0.0075					0.0625			0.0625	0.3675	0.0500	0.0050			
9.1									0.0025											
9.2															0.0025					
9.3											0.1950									
10			0.2575	0.2375	0.0300	0.0900		0.0050	0.5875					0.0475	0.1100	0.2550		0.0050		
11	0.0025		0.1925	0.2650	0.2350	0.0525		0.0050	0.2025	0.0050			0.0450	0.5025	0.1300	0.3675	0.0025	0.0125		
11.2										0.0025										
11.3									0.0325											
12	0.0025		0.2400	0.4025	0.1075	0.1375		0.1250	0.0425	0.0200			0.0025	0.1825	0.1800	0.2475	0.0350	0.0275		
12.2										0.0125										
12.3				0.0025																
13	0.0050	0.0025	0.0800	0.0650	0.0050	0.3400		0.1525	0.0125	0.1675				0.0850	0.0925	0.0300	0.2200	0.1400		
13.2										0.1200										
14	0.0425	0.0275	0.0175	0.0050		0.2100		0.2325	0.1050	0.3300			0.0125	0.0050	0.0675	0.0050	0.4275	0.1250	0.0025	
14.1										0.0025										
14.2										0.0325										
15	0.4950	0.0500	0.0025	0.0025		0.1325		0.1500	0.0125	0.1725			0.4650		0.0025		0.2500	0.1450	0.0050	
15.2										0.1000										

15.3																		0.0150		
16	0.2925	0.3450				0.0225		0.1200	0.0025	0.0225			0.4125				0.0375	0.2175	0.0125	0.0150
16.2										0.0075										
16.3																		0.0425		
17	0.1125	0.3775				0.0075		0.1375		0.0050			0.0500		0.0025		0.0200	0.0425	0.0625	0.1800
17.2																				
17.3																		0.1850	0.0100	
18	0.0400	0.1425						0.0325				0.0125	0.0100				0.0025	0.0025	0.1850	0.0675
18.3																		0.0350	0.0050	
19	0.0075	0.0450						0.0175				0.1125	0.0025						0.2725	0.2250
19.1																			0.0025	
19.3																		0.0050	0.0075	
20		0.0100						0.0125				0.0675							0.2650	0.1450
20.3																			0.0050	
21								0.0025				0.0925							0.0650	0.0225
21.3																				
22								0.0050				0.0525							0.0400	0.0775
22.2												0.0025								
22.3																				
23												0.0800							0.0400	0.1800
24												0.1825							0.0100	0.0500
24.3																				
25								0.0025				0.1825							0.0100	0.0300
25.3																				



26												0.1700								0.0075
27							0.0050					0.0425								
28							0.0675					0.0025								
29							0.1575													
29.2							0.0025													
30							0.2050													
30.2							0.0350													
31							0.0575													
31.2							0.1350													
32							0.0025													
32.2							0.2300													
33.2							0.0875													
34.2							0.0125													
35							0.0025													
N	200	200	200	200	200	200	200	200	200	200	200	200	200	200	200	200	200	200	200	200
Frecuencia mínima	0.0125	0.0125	0.0125	0.0125	0.0125	0.0125	0.0125	0.0125	0.0125	0.0125	0.0125	0.0125	0.0125	0.0125	0.0125	0.0125	0.0125	0.0125	0.0125	0.0125
HW	0.8403	0.7410	0.4577	0.1487	0.7014	0.1291	0.2965	0.8443	0.9428	0.0382	0.0525	0.4701	0.2209	0.8979	0.4543	0.1945	0.9813	0.1813	0.7518	0.7137
Het(esp)	0.6549	0.7146	0.7912	0.7087	0.5906	0.7944	0.8472	0.8518	0.6011	0.8087	0.6930	0.8698	0.6104	0.6904	0.7904	0.7353	0.7050	0.8583	0.8114	0.8509
Het(obs)	0.6800	0.7000	0.8200	0.6800	0.5700	0.7800	0.7900	0.8400	0.6200	0.8300	0.6150	0.8250	0.6300	0.7100	0.7950	0.7400	0.6950	0.8550	0.8200	0.8250

Siendo N: el número de muestras, H-W: p-value para el equilibrio de Hardy Weinberg, Het(esp): Heterocigosidad esperada, Het(obs): Heterocigosidad observada

## 5.2. Análisis de los parámetros poblacionales

De los 20 locis analizados (tabla 2) 13 presentaron un alto nivel de Heterocigosidad observada (Het (obs)) con un rango de 0.57 (TPOX) a 0.855 (D1S1656), siendo los de más alto valor D19S433 (0.8300), D18S51 (0.8400), D1S1656 (0.8550) y los de menor valor TPOX (0.57), TH01 (0.6150), D2S441 (0.6200). En cuanto a la Heterocigosidad esperada (Het (esp)) esta se encontró en el rango de 0.5906 (TPOX) a 0.8698 (FGA), siendo los de mayor valor D18S51 (0.8518), D1S1656 (0.8583), FGA (0.8698) y los de menor valor TPOX (0.5906), D2S441 (0.6011), D22S1045 (0.6104). De otro lado, los locis más polimórficos fueron D18S51, D19S433, D1S1656, con 14 alelos cada uno y D12S391 con 17 alelos, y el menos polimórfico fue el TH01 con 5 alelos. El loci D2S441 (alelo 10) fue el que presentó el más alto nivel de frecuencia alélica con un valor de 0.5875.

De los 20 marcadores STRs analizados, 19 se encontraron en equilibrio de Hardy- Weinberg, a excepción del locus D19S433 (0.0382), para lo cual se utilizó la corrección de Bonferroni, asumiendo un nivel de significancia de  $p > 0.0025$ .

La frecuencia mínima para los 20 marcadores analizados fue de 0.0125.

La prueba de desequilibrio de ligamiento fue realizado para todos los posibles pares de combinaciones entre los 20 locis indicados en la tabla 3, los resultados muestran que 11 pares (D3S1358- D10S1248, vWA- D2S1338, D16S539- D7S820, CSF1PO- TH01, TPOX- D19S433, D21S11- D13S317, D21S11- D7S820, D21S11- D12S391, D18S51- TH01, D18S51- D1S1656, D13S317- D2S1338) de las 190 comparaciones entre pares evidencian desequilibrio de ligamiento ( $p < 0.05$ ), sin embargo ningún desequilibrio de

ligamiento permaneció luego de aplicar la corrección de Bonferroni (a un nivel de significancia  $p > 0.000263$ ).

Se encontraron alelos que todavía no han sido descritos en la base de datos del THE NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY (NIST) : 12.3 del marcador CSF1PO, 8.2 del marcador TPOX, 9.1 del marcador D2S441, 11.2 del marcador D19S433, 9.2 del marcador D13S317 y 14, 19.1, 20.3 del marcador D12S391.

**Tabla 3.** Valores p para el análisis de Desequilibrio de Ligamiento.

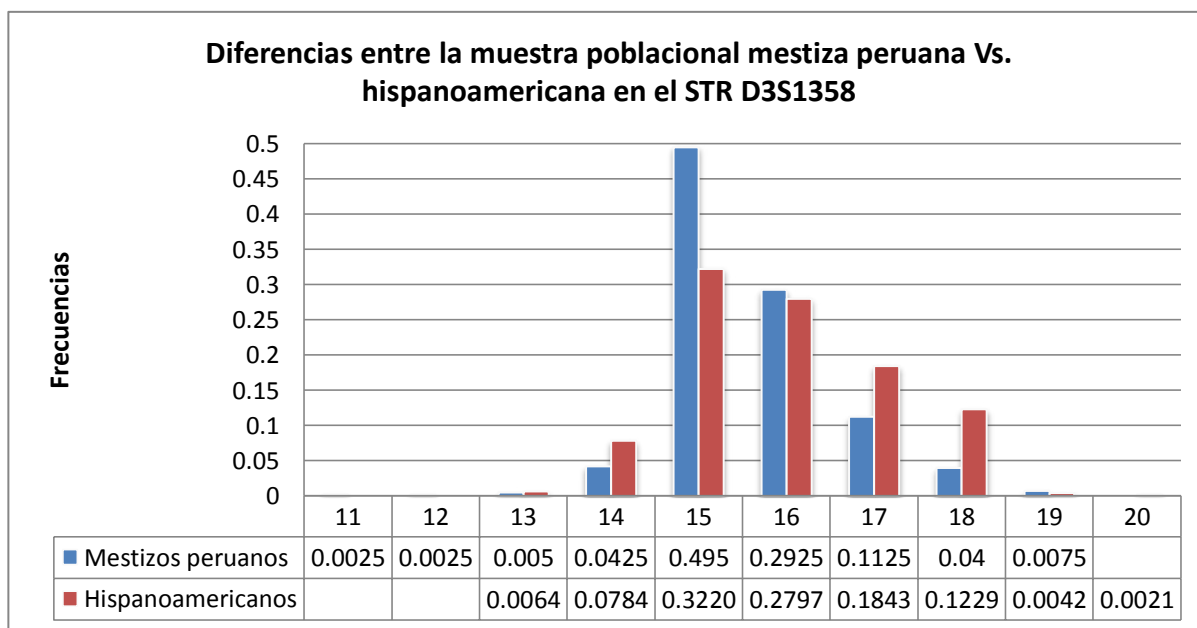
Par de loci	p-value	Par de loci	p-value	Par de loci	p-value	Par de loci	p-value
Par(0,1)	0.25812	Par(3,10)	0.04327	Par(5,14)	0.48851	Par(8,17)	0.20832
Par(0,2)	0.55554	Par(4,10)	0.40129	Par(6,14)	0.03980	Par(9,17)	0.07495
Par(1,2)	0.68455	Par(5,10)	0.87198	Par(7,14)	0.88139	Par(10, 17)	0.91505
Par(0,3)	0.75129	Par(6,10)	0.30277	Par(8,14)	0.16762	Par(11, 17)	0.11941
Par(1,3)	0.71901	Par(7,10)	0.04317	Par(9,14)	0.96604	Par(12, 17)	0.49059
Par(2,3)	0.07446	Par(8,10)	0.10099	Par(10, 14)	0.56901	Par(13, 17)	0.20287
Par(0,4)	0.11089	Par(9,10)	0.46990	Par(11, 14)	0.40139	Par(14, 17)	0.24109
Par(1,4)	0.53465	Par(0,11)	0.85149	Par(12, 14)	0.76168	Par(15, 17)	0.34089
Par(2,4)	0.64287	Par(1,11)	0.61990	Par(13, 14)	0.82960	Par(16, 17)	0.36663
Par(3,4)	0.79564	Par(2,11)	0.59624	Par(0,15)	0.93772	Par(0,18)	0.30149
Par(0,5)	0.49089	Par(3,11)	0.64485	Par(1,15)	0.98238	Par(1,18)	0.21149
Par(1,5)	0.61960	Par(4,11)	0.85921	Par(2,15)	0.02386	Par(2,18)	0.30020
Par(2,5)	0.09327	Par(5,11)	0.33703	Par(3,15)	0.16178	Par(3,18)	0.38693
Par(3,5)	0.45465	Par(6,11)	0.06921	Par(4,15)	0.25921	Par(4,18)	0.14723
Par(4,5)	0.58802	Par(7,11)	0.09050	Par(5,15)	0.85564	Par(5,18)	0.31277
Par(0,6)	0.92465	Par(8,11)	0.22000	Par(6,15)	0.00970	Par(6,18)	0.01594
Par(1,6)	0.08416	Par(9,11)	0.17644	Par(7,15)	0.30119	Par(7,18)	0.96970
Par(2,6)	0.66208	Par(10, 11)	0.74505	Par(8,15)	0.52782	Par(8,18)	0.50535
Par(3,6)	0.05198	Par(0,12)	0.22465	Par(9,15)	0.57525	Par(9,18)	0.82347
Par(4,6)	0.95812	Par(1,12)	0.22446	Par(10, 15)	0.45198	Par(10, 18)	0.22881
Par(5,6)	0.77624	Par(2,12)	0.76921	Par(11, 15)	0.44396	Par(11, 18)	0.22297
Par(0,7)	0.98277	Par(3,12)	0.58307	Par(12, 15)	0.43792	Par(12, 18)	0.80386
Par(1,7)	0.44723	Par(4,12)	0.33545	Par(13, 15)	0.80208	Par(13, 18)	0.23931
Par(2,7)	0.97901	Par(5,12)	0.45465	Par(14, 15)	0.98644	Par(14, 18)	0.31634
Par(3,7)	0.58515	Par(6,12)	0.35020	Par(0,16)	0.00030	Par(15, 18)	0.87089
Par(4,7)	0.84842	Par(7,12)	0.61842	Par(1,16)	0.42277	Par(16, 18)	0.80455
Par(5,7)	0.56921	Par(8,12)	0.39525	Par(2,16)	0.21010	Par(17, 18)	0.77713
Par(6,7)	0.59327	Par(9,12)	0.45970	Par(3,16)	0.99683	Par(0,19)	0.77386
Par(0,8)	0.06248	Par(10, 12)	0.32396	Par(4,16)	0.55644	Par(1,19)	0.03861
Par(1,8)	0.52178	Par(11, 12)	0.78673	Par(5,16)	0.65069	Par(2,19)	0.62109
Par(2,8)	0.50376	Par(0,13)	0.28604	Par(6,16)	0.54832	Par(3,19)	0.40356
Par(3,8)	0.13267	Par(1,13)	0.96030	Par(7,16)	0.56960	Par(4,19)	0.57188
Par(4,8)	0.14653	Par(2,13)	0.20653	Par(8,16)	0.14139	Par(5,19)	0.17178
Par(5,8)	0.24594	Par(3,13)	0.60772	Par(9,16)	0.11752	Par(6,19)	0.80812
Par(6,8)	0.59149	Par(4,13)	0.25178	Par(10, 16)	0.26376	Par(7,19)	0.86713
Par(7,8)	0.49297	Par(5,13)	0.28050	Par(11, 16)	0.14069	Par(8,19)	0.60743
Par(0,9)	0.86149	Par(6,13)	0.27020	Par(12, 16)	0.55495	Par(9,19)	0.11782
Par(1,9)	0.74931	Par(7,13)	0.57871	Par(13, 16)	0.65644	Par(10, 19)	0.85416
Par(2,9)	0.53703	Par(8,13)	0.31881	Par(14, 16)	0.62149	Par(11, 19)	0.23842
Par(3,9)	0.83208	Par(9,13)	0.23554	Par(15, 16)	0.52406	Par(12, 19)	0.29525
Par(4,9)	0.02109	Par(10, 13)	0.70762	Par(0,17)	0.20911	Par(13, 19)	0.62901
Par(5,9)	0.72743	Par(11, 13)	0.14554	Par(1,17)	0.93178	Par(14, 19)	0.02970
Par(6,9)	0.73297	Par(12, 13)	0.80525	Par(2,17)	0.32703	Par(15, 19)	0.66158
Par(7,9)	0.34673	Par(0,14)	0.77317	Par(3,17)	0.25267	Par(16, 19)	0.17574
Par(8,9)	0.53406	Par(1,14)	0.75564	Par(4,17)	0.62257	Par(17, 19)	0.98950
Par(0,10)	0.08861	Par(2,14)	0.09881	Par(5,17)	0.46861	Par(18, 19)	0.22762
Par(1,10)	0.11505	Par(3,14)	0.57455	Par(6,17)	0.71119		
Par(2,10)	0.57535	Par(4,14)	0.69931	Par(7,17)	0.01772		

Nominaciones de loci: D3S1358 = locus 0, vWA = locus 1, D16S539 = locus 2, CSF1PO = locus 3, TPOX = locus 4, D8S1179 = locus 5, D21S11 = locus 6, D18S51 = locus 7, D2S441 = locus 8, D19S433 = locus 9, TH01 = locus 10, FGA = locus 11, D22S1045 = locus 12, D5S818 = locus 13, D13S317 = locus 14, D7S820 = locus 15, D10S1248 = locus 16, D1S1656 = locus 17, D12S391 = locus 18, D2S1338 = locus 19.

### 5.3. Variabilidad de las frecuencias alélicas de la muestra poblacional mestiza peruana e hispanoamericana.

#### 5.3.1. Marcador D3S1358

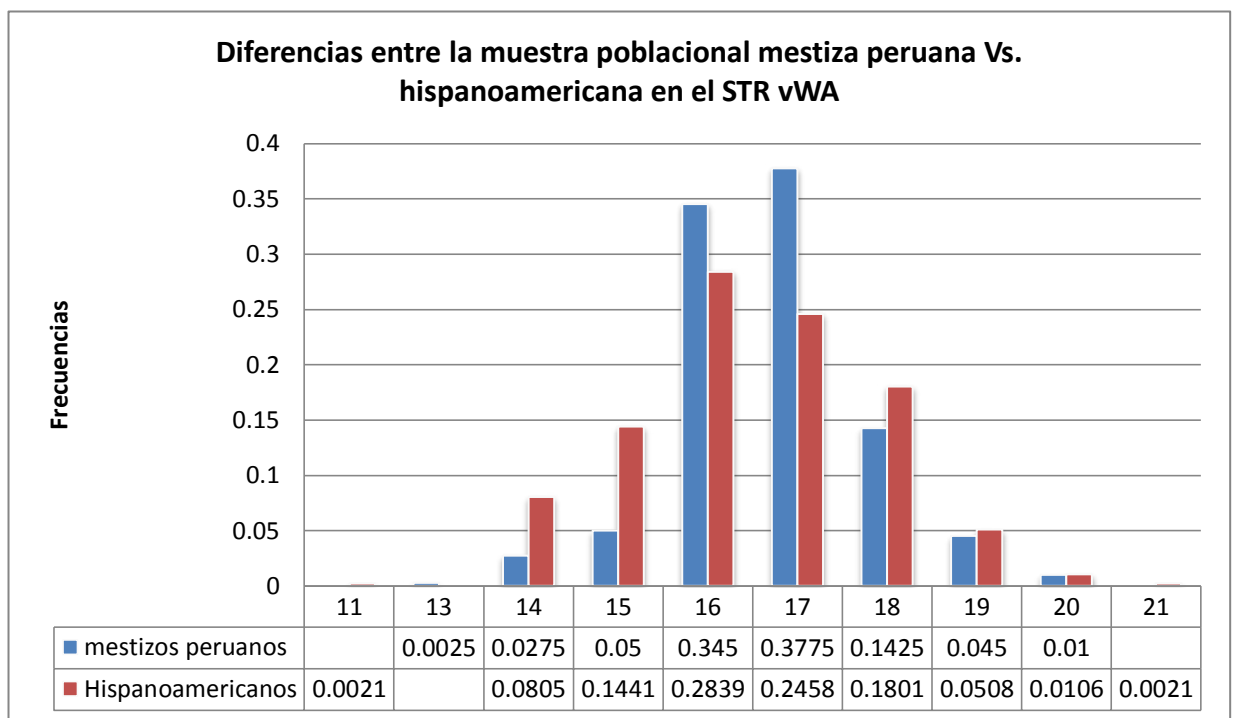
La muestra poblacional peruana para este marcador posee los alelos 11 y 12 (figura 5) que no posee la población hispanoamericana de Hill *et al.* (2013). Asimismo, el alelo 20 está presente solo en la población hispanoamericana. El alelo con mayor frecuencia en la muestra poblacional peruana trabajada es el alelo 15 con 0.495, siendo también este alelo el de mayor frecuencia en la población hispanoamericana. Los alelos de menor frecuencia en la muestra poblacional peruana fueron el 11 y 12, mientras que en la población hispanoamericana fue el 20.



**Figura 5.** Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR D3S1358.

### 5.3.2. Marcador vWA

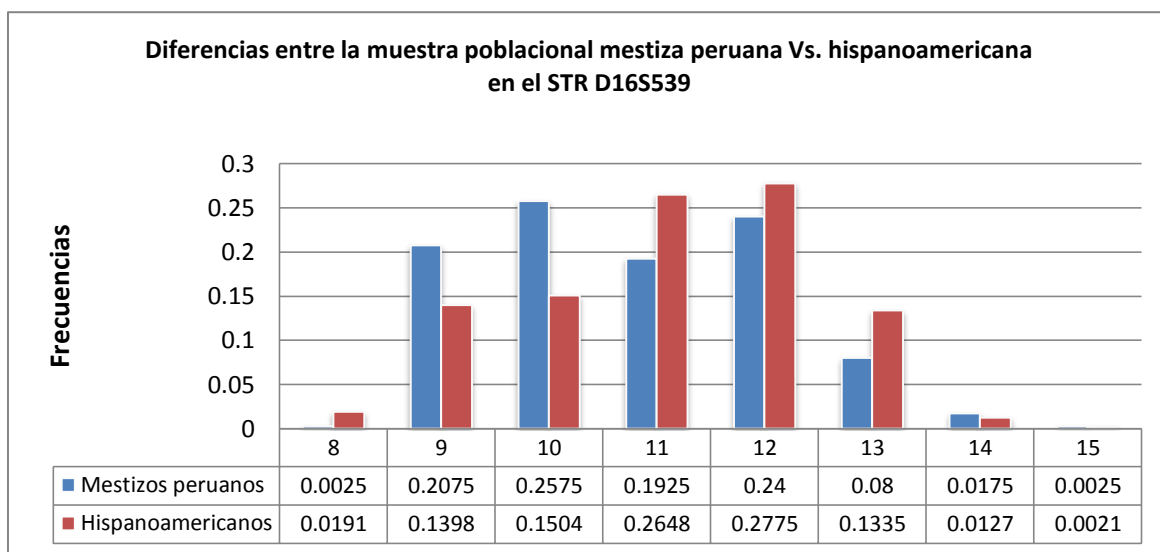
La muestra poblacional analizada en este trabajo, para este marcador, posee el alelo 13 (figura 6) que no posee la población hispanoamericana de Hill *et al.* (2013). Además, la población hispanoamericana presenta los alelos 11 y 21 que no presenta la muestra poblacional peruana. El alelo de mayor frecuencia en la muestra poblacional peruana fue el alelo 17 con 0.3775, mientras que en la población hispanoamericana fue el alelo 16 con 0.2839. El alelo con menor frecuencia en la muestra poblacional peruana fue alelo 13 y en la población hispanoamericana fue el alelo 11 y 21.



**Figura 6.** Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR vWA.

### 5.3.3. Marcador D16S539

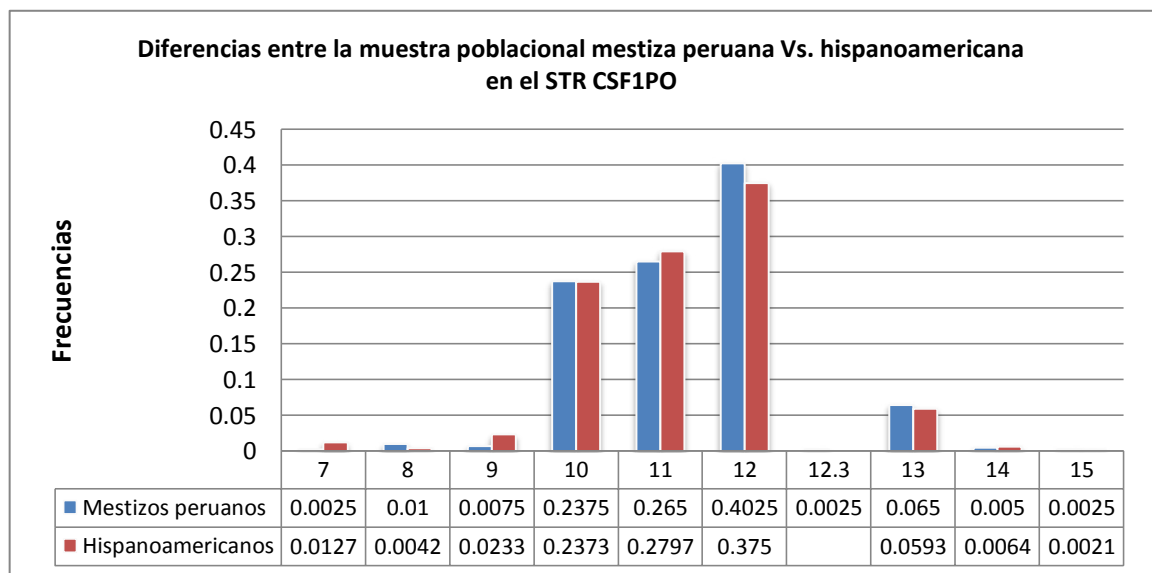
El alelo de mayor frecuencia en la muestra poblacional peruana es el alelo 10 (figura 7) con 0.2575, mientras que en la población hispanoamericana fue el alelo 12 con 0.2775, por otro lado, los alelos con menor frecuencia en la muestra poblacional peruana fueron los alelos 8 y 15, mientras que en la población hispanoamericana solo fue el alelo 15.



**Figura 7.** Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR D16S539.

#### 5.3.4. Marcador CSF1PO

La muestra poblacional peruana para este marcador presenta el alelo 12.3 (figura 8) que no se encuentra en la población hispanoamericana de Hill *et al.* (2013). El alelo de mayor frecuencia en la muestra poblacional peruana es el alelo 12 con 0.4025, mientras que en la población hispanoamericana el alelo de mayor frecuencia también fue el alelo 12. Los alelos de menor frecuencia en la muestra poblacional peruana fueron los alelos 7, 12.3, y 15, mientras que en la población hispanoamericana solo fue el alelo 15.

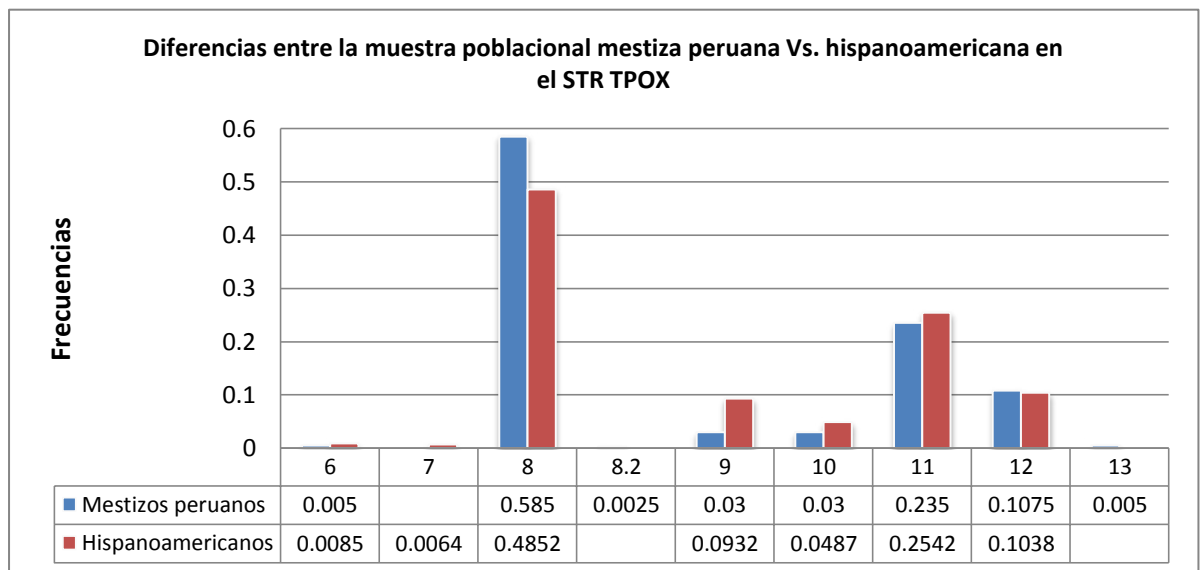


**Figura 8.** Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR CSF1PO.



### 5.3.5. Marcador TPOX

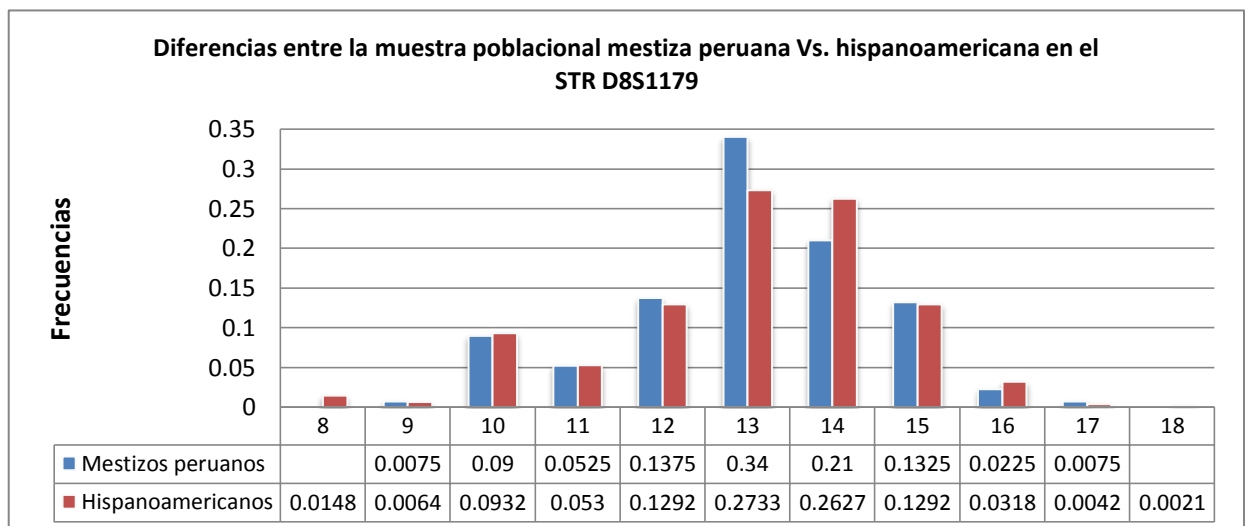
La muestra poblacional peruana para este marcador presenta los alelos 8.2 y 13 (figura 9) que no posee la población hispanoamericana de Hill *et al.* (2013), además la población hispanoamericana presenta el alelo 7 que no presenta la muestra poblacional peruana. El alelo de mayor frecuencia en la muestra poblacional peruana es el alelo 8 con 0.585, siendo también este alelo el de mayor frecuencia en la población hispanoamericana con 0.4852, por otro lado, el alelo de menor frecuencia en la muestra poblacional peruana fue el alelo 8.2 y en la población hispanoamericana fue el alelo 7.



**Figura 9.** Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR TPOX.

### 5.3.6. Marcador D8S1179

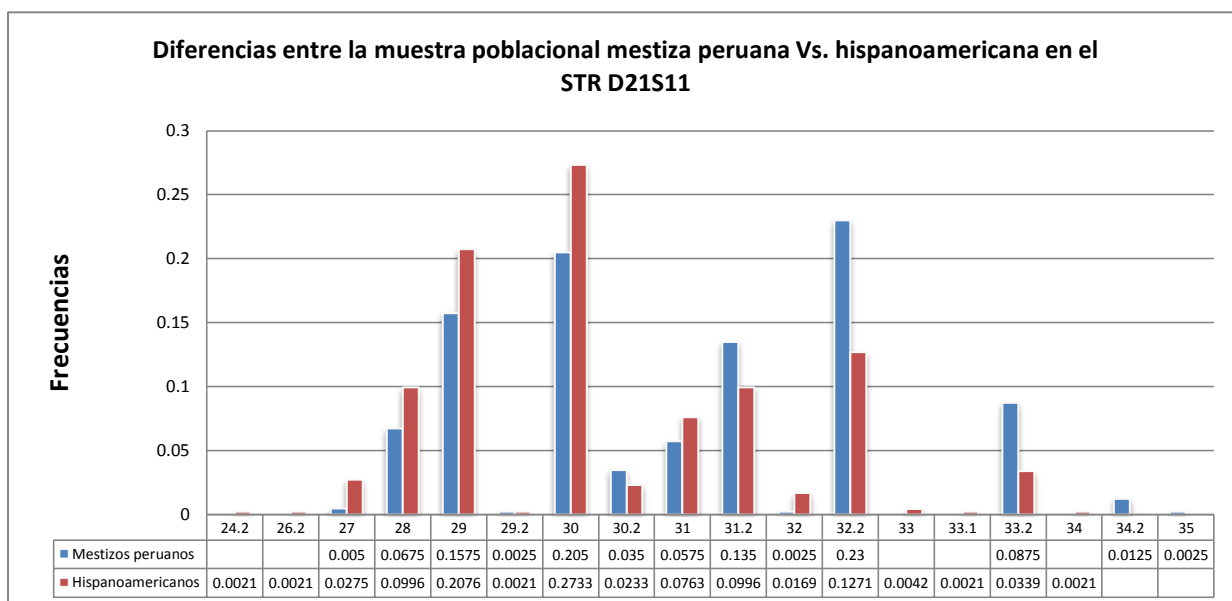
La población hispanoamericana para este marcador presenta los alelos 8 y 18 (figura 10) que no están presentes en la muestra poblacional peruana. El alelo con mayor frecuencia en la muestra poblacional peruana fue el alelo 12 con 0.1375, y en la población hispanoamericana fue el alelo 13 con 0.2733, en cambio los alelos de menor frecuencia en la muestra poblacional peruana fueron 9 y 17, mientras que en la población hispanoamericana fue el alelo 18.



**Figura 10.** Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR D8S1179.

### 5.3.7. Marcador D21S11

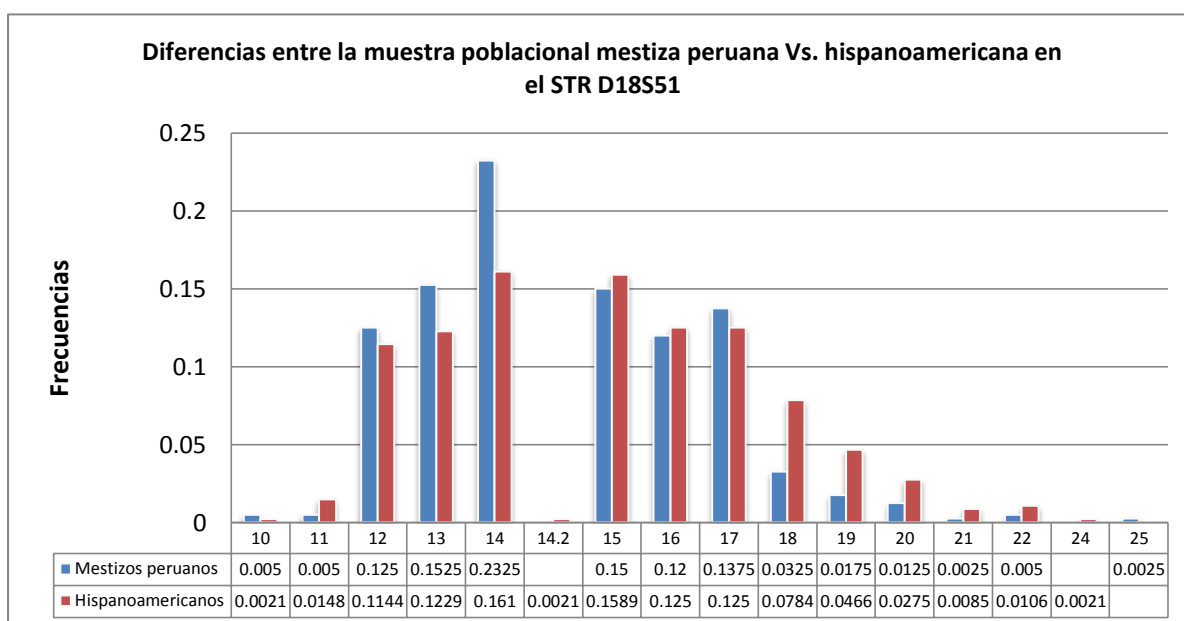
La muestra poblacional peruana para este marcador presenta los alelos 34.2 y 35 (figura 11) que no posee la población hispanoamericana de Hill *et al.* (2013), además la población hispanoamericana presenta los alelos 24.2, 26.2, 33, 33.1 y 34 que no presenta la muestra poblacional peruana. El alelo de mayor frecuencia en la muestra poblacional peruana fue el alelo 30 con 0.205, mientras que en la población hispanoamericana también fue este alelo con 0.2733, por otro lado, los alelos con menor frecuencia en la muestra poblacional peruana fueron 29.2, 32 y 35, mientras que en la población hispanoamericana fueron los alelos 24.2, 26.2, 29.2, 33.1 y 34.



**Figura 11.** Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR D21S11.

### 5.3.8. Marcador D18S51

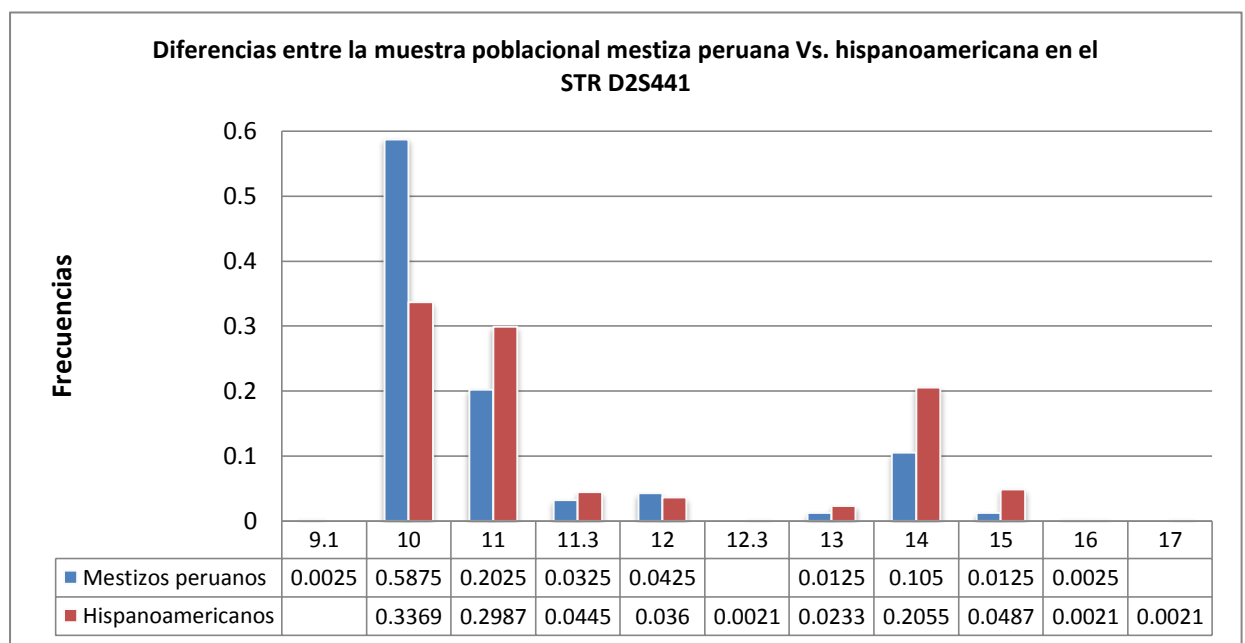
La muestra poblacional peruana para este marcador presenta el alelo 25 (figura 12) que no posee la población hispanoamericana de Hill *et al.* (2013), además la población hispanoamericana presenta los alelos 14.2, 24 que no presenta la muestra poblacional peruana. El alelo de mayor frecuencia en la muestra poblacional peruana fue el alelo 14 con 0.2325, siendo también este alelo el de mayor frecuencia en la población hispanoamericana con 0.161, por otro lado, los alelos de menor frecuencia en la muestra poblacional peruana fueron el 21 y 25, mientras que en la población hispanoamericana fueron el 10, 14.2, y 24.



**Figura 12.** Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR D18S51.

### 5.3.9. Marcador D2S441

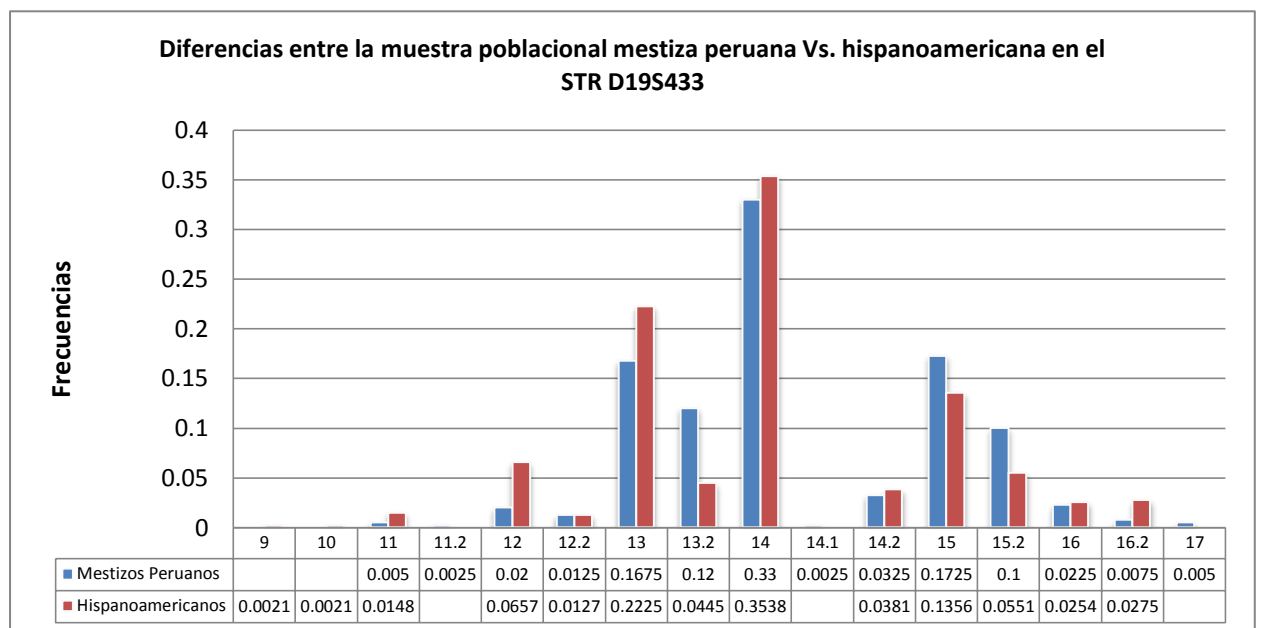
La muestra poblacional peruana para este marcador presenta el alelo 9.1 (figura 13) que no posee la población hispanoamericana de Hill *et al.* (2013), además la población hispanoamericana presenta los alelos 12.3 y 17 que no presenta la muestra poblacional peruana. El alelo de mayor frecuencia en la muestra poblacional peruana fue el alelo 10 con 0.5875, siendo este también el alelo de mayor frecuencia en la población hispanoamericana con 0.3369, por otro lado, los alelos de menor frecuencia en la muestra poblacional peruana fueron el 9.1 y 16, mientras que en la población hispanoamericana fueron el 12.3, 16, y 17.



**Figura 13.** Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR D2S441.

### 5.3.10. Marcador D19S433

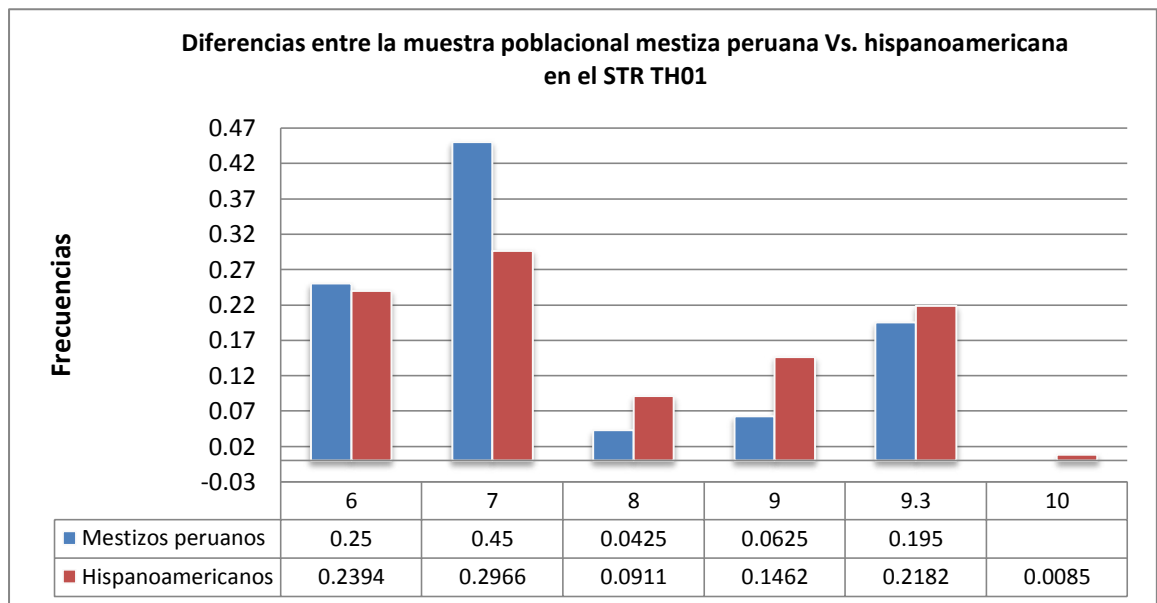
La muestra poblacional peruana para este marcador presenta los alelos 11.2, 14.1 y 17 (figura 14) que no posee la población hispanoamericana de Hill *et al.* (2013), además la población hispanoamericana presenta los alelos 9, 10 que no presenta la muestra poblacional peruana. El alelo de mayor frecuencia en la muestra poblacional peruana fue el alelo 14 con 0.33, siendo este también el alelo de mayor frecuencia en la población hispanoamericana con 0.3538, por otro lado, los alelos de menor frecuencia en la muestra poblacional peruana fueron el 11.2 y 14.1, mientras que en la población hispanoamericana fueron el 9, 10.



**Figura 14.** Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR D19S433.

### 5.3.11. Marcador TH01

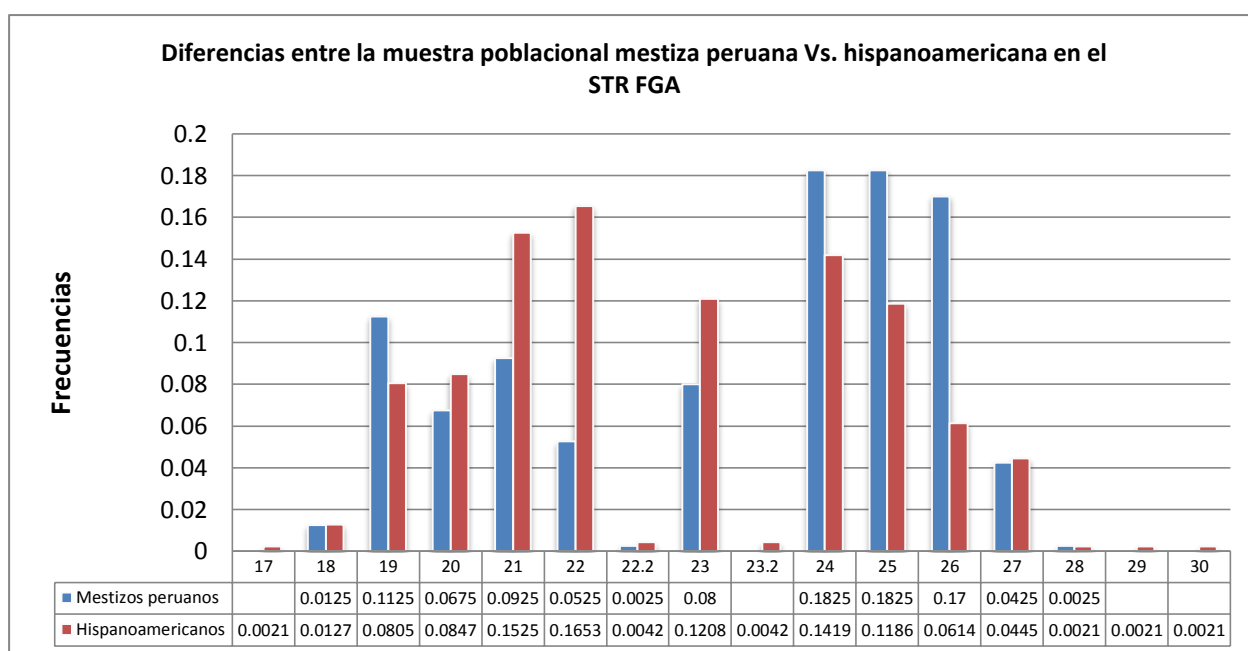
La población hispanoamericana para este marcador presenta el alelo 10 (figura 15) que no está presente en la muestra poblacional peruana. El alelo con mayor frecuencia en la muestra poblacional peruana fue el alelo 7 con 0.45, y en la población hispanoamericana también fue este alelo con una frecuencia de 0.2966, en cambio el alelo de menor frecuencia en la muestra poblacional peruana fue el 9, mientras que en la población hispanoamericana fue el alelo 10.



**Figura 15.** Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR TH01.

### 5.3.12. Marcador FGA

La población hispanoamericana para este marcador presenta el alelo 17, 23.2, 29 y 30 (figura 16) que no está presente en la muestra poblacional peruana. Los alelos con mayor frecuencia en la muestra poblacional peruana fueron el 24 y 25 con 0.1825 cada uno, mientras que en la población hispanoamericana fue el alelo 22 con 0.1653, en cambio los alelos con menor frecuencia en la muestra poblacional peruana fueron el 22.2 y 28, mientras que en la población hispanoamericana fueron el 17, 28, 29, 30.

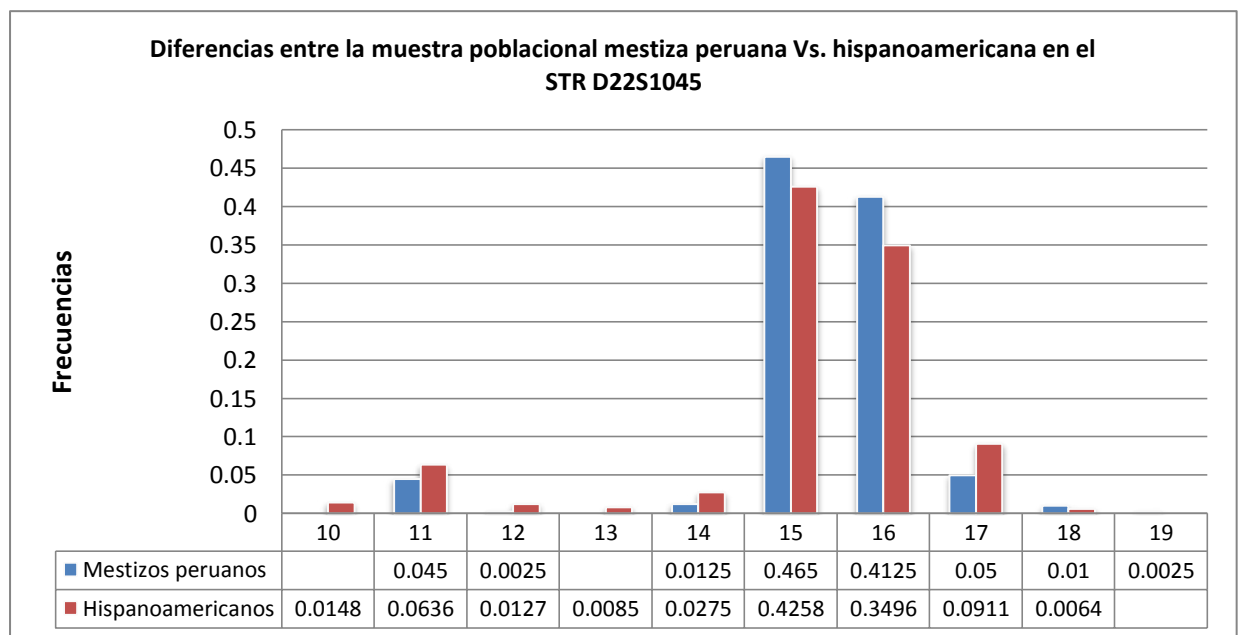


**Figura 16.** Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR FGA.



### 5.3.13. Marcador D22S1045

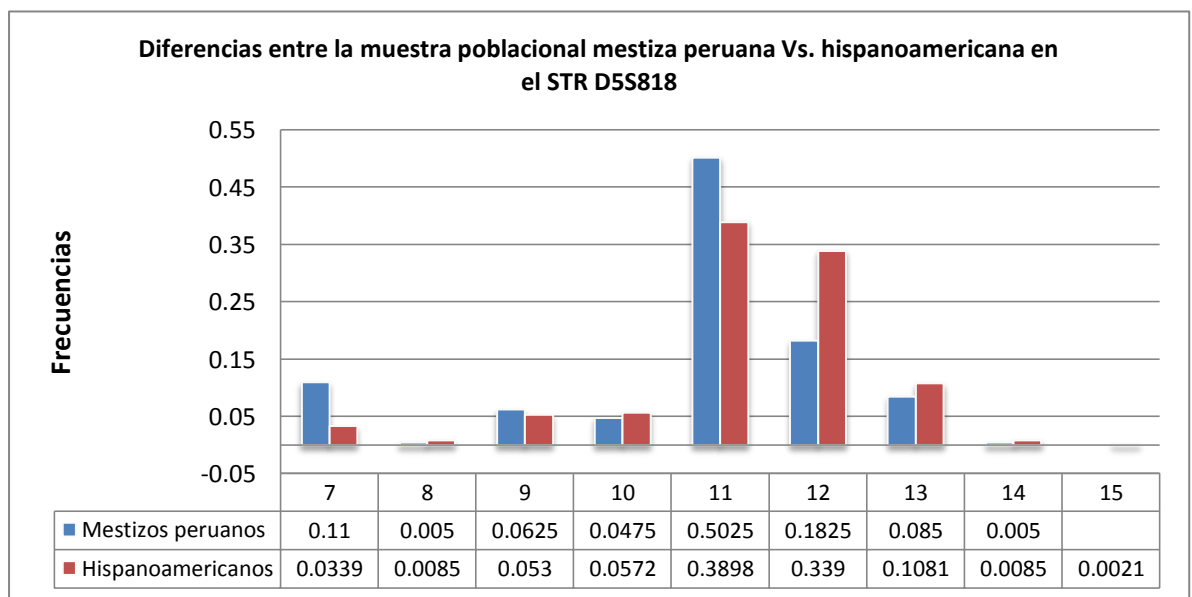
La muestra poblacional peruana para este marcador presenta el alelo 19 (figura 17) que no posee la población hispanoamericana de Hill *et al.* (2013), además la población hispanoamericana presenta los alelos 10 y 13 que no presenta la muestra poblacional peruana. El alelo de mayor frecuencia en la muestra poblacional peruana fue el alelo 15 con 0.465, y en la población hispanoamericana también fue este alelo con una frecuencia de 0.4258, por otro lado, los alelos de menor frecuencia alélica en la muestra poblacional peruana fueron los alelos 12 y 19, mientras que en la población hispanoamericana fue el alelo 18.



**Figura 17.** Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR D22S1045.

### 5.3.14. Marcador D5S818

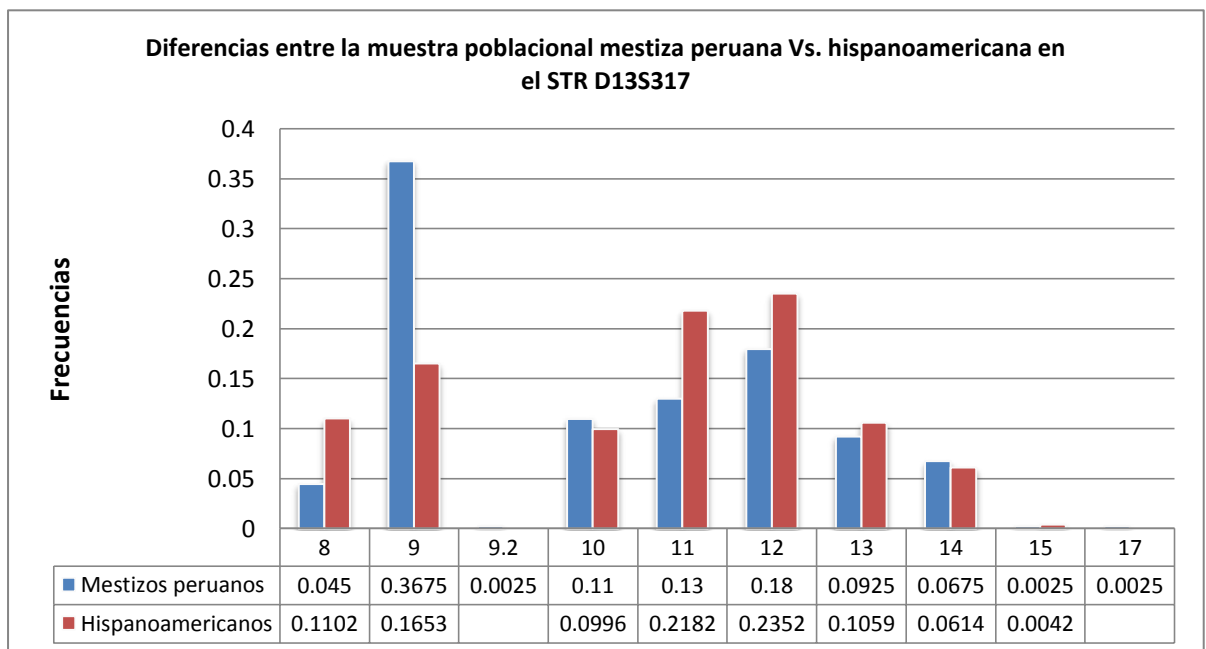
La población hispanoamericana para este marcador presenta el alelo 15 (figura 18) que no se encuentra presente en la muestra poblacional peruana. El alelo con mayor frecuencia en la muestra poblacional peruana fue el 11 con 0.5025, y en la población hispanoamericana también fue este alelo con una frecuencia de 0.3898, por otro lado, los alelos de menor frecuencia alélica en la muestra poblacional peruana fueron los alelos 8 y 14, mientras que en la población hispanoamericana fue el alelo 15.



**Figura 18.** Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR D5S818.

### 5.3.15. Marcador D13S317

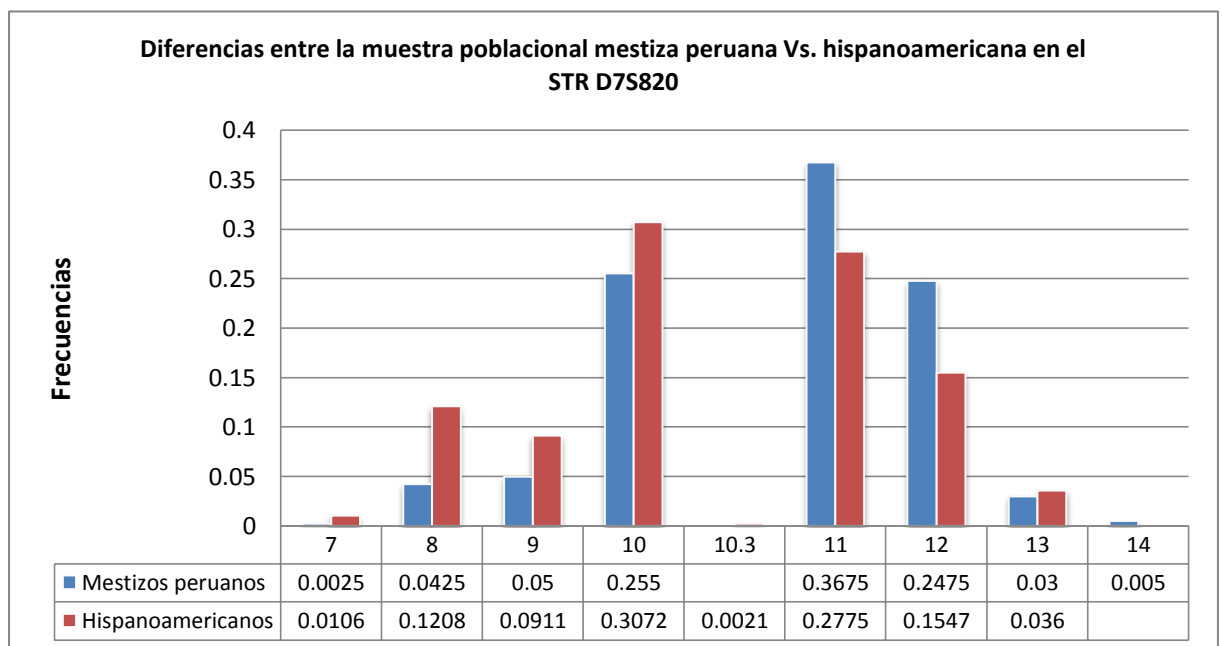
La muestra poblacional peruana para este marcador presenta los alelos 9.2 y 17 (figura 19) que no posee la población hispanoamericana de Hill *et al.* (2013). El alelo de mayor frecuencia en la muestra poblacional peruana fue el alelo 9 con 0.3675, y en la población hispanoamericana fue el alelo 12 con 0.2352, por otro lado, los alelos de menor frecuencia alélica en la muestra poblacional peruana fueron 9.2, 15 y 17, mientras que en la población hispanoamericana fue el alelo 15.



**Figura 19.** Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR D13S317.

### 5.3.16. Marcador D7S820

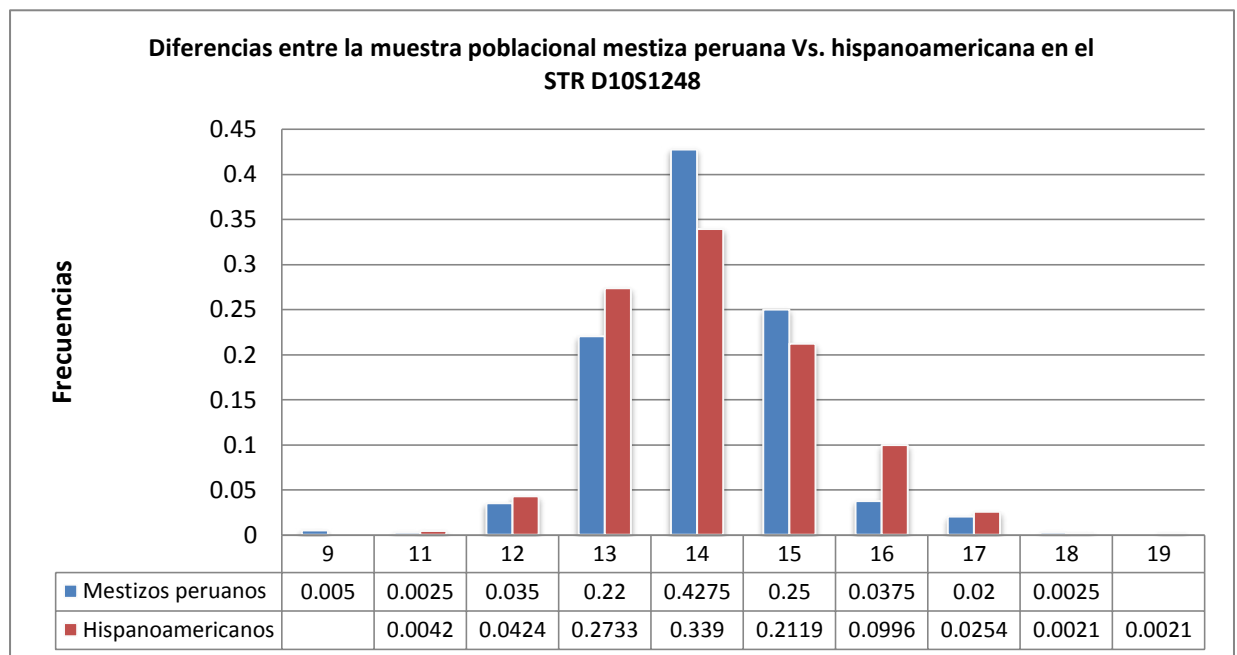
La muestra poblacional peruana para este marcador presenta el alelo 14 (figura 20) que no posee la población hispanoamericana de Hill *et al.* (2013), además la población hispanoamericana presenta los alelos 10.3 y 14 que no presenta la muestra poblacional peruana. El alelo de mayor frecuencia en la muestra poblacional peruana fue el alelo 11, mientras que en la población hispanoamericana fue el alelo 10, por otro lado, el alelo de menor frecuencia alélica en la muestra poblacional peruana fue el alelo 7, mientras que en la población hispanoamericana fue el alelo 10.3.



**Figura 20.** Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR D7S820.

### 5.3.17. Marcador D10S1248

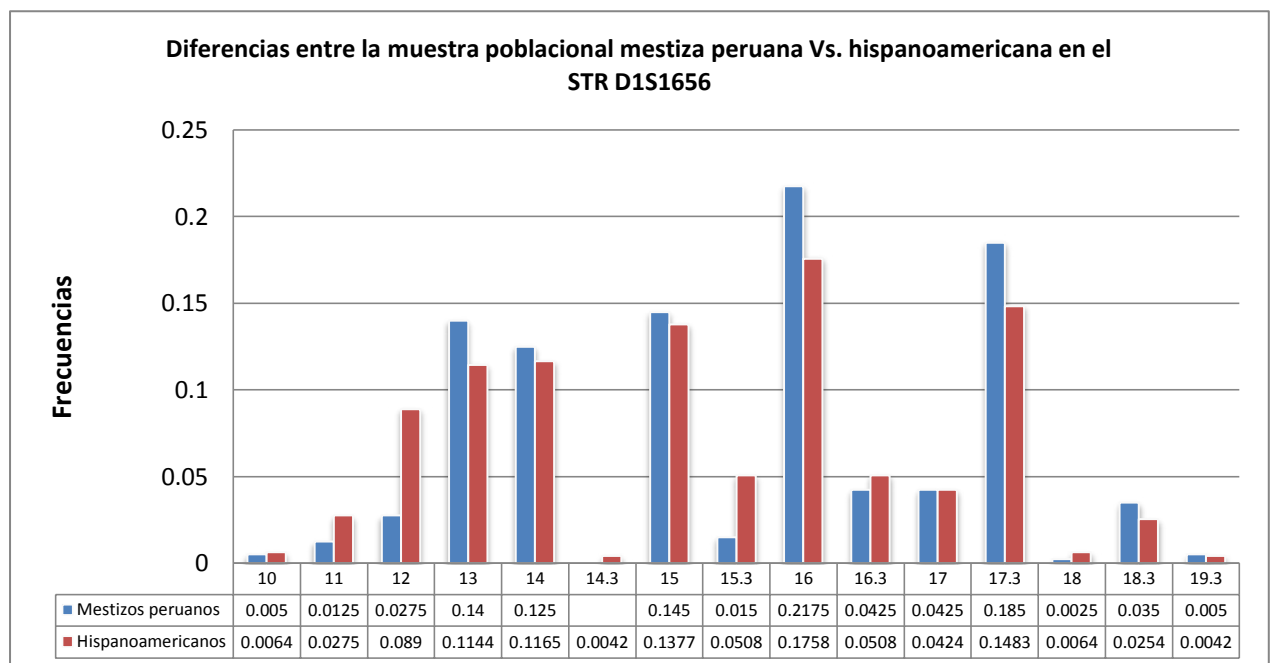
La muestra poblacional peruana para este marcador presenta el alelo 9 (figura 21) que no posee la población hispanoamericana de Hill *et al.* (2013), además la población hispanoamericana presenta el alelo 19 que no presenta la muestra poblacional peruana. El alelo de mayor frecuencia en la muestra poblacional peruana fue el 14 y en la población hispanoamericana también fue este alelo con una frecuencia de 0.339, por otro lado, los alelos de menor frecuencia alélica en la muestra poblacional peruana fueron el 11 y 18, mientras que en la población hispanoamericana fueron el 18 y 19.



**Figura 21.** Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR D10S1248.

### 5.3.18. Marcador D1S1656

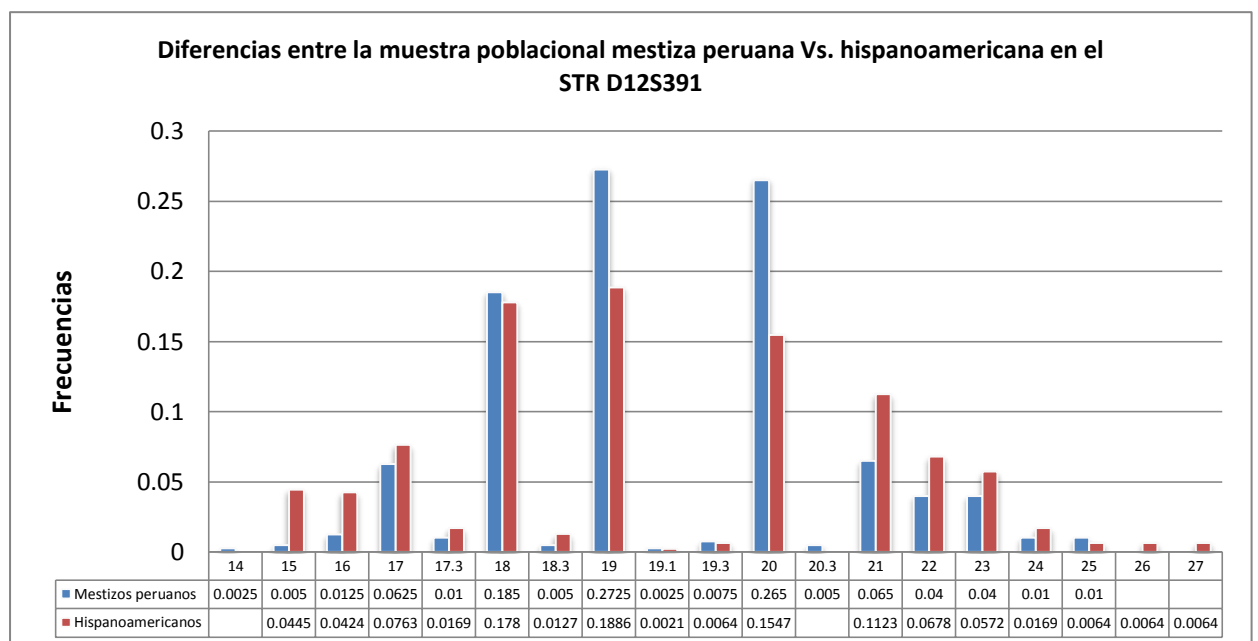
La población hispanoamericana para este marcador presenta el alelo 14.3 (figura 22) que no se encuentra presente en la muestra poblacional peruana. El alelo con mayor frecuencia en la muestra poblacional peruana fue el alelo 16 con 0.2175, y en la población hispanoamericana también fue este alelo con una frecuencia de 0.1758, por otro lado, el alelo de menor frecuencia alélica en la muestra poblacional peruana fue el 18, mientras que en la población hispanoamericana fueron el 14.3 y 19.3.



**Figura 22.** Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR D1S1656.

### 5.3.19. Marcador D12S391

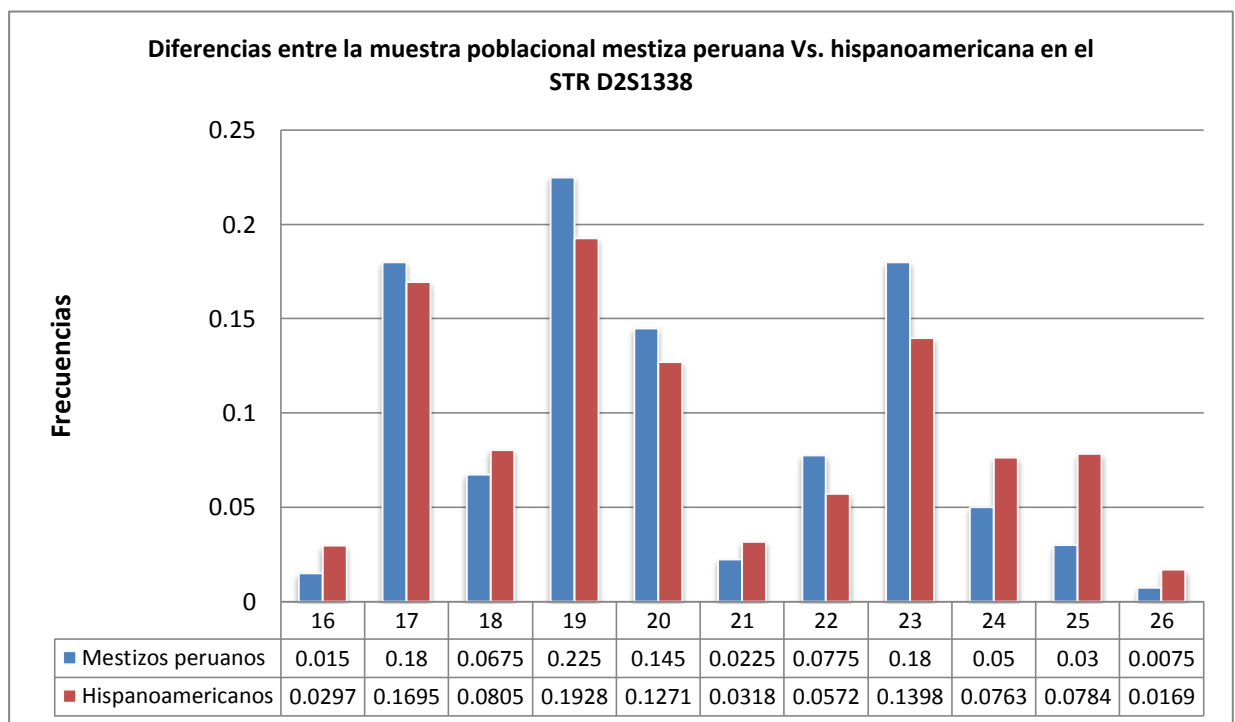
La muestra poblacional peruana para este marcador presenta los alelos 14 y 20.3 (figura 23) que no posee la población hispanoamericana de Hill *et al.* (2013), además la población hispanoamericana presenta los alelos 26 y 27 que no presenta la muestra poblacional peruana. El alelo de mayor frecuencia en la muestra poblacional peruana fue el alelo 19 con 0.2725, y en la población hispanoamericana también fue este alelo con una frecuencia de 0.1886, por otro lado, los alelos de menor frecuencia alélica en la muestra poblacional peruana fueron el 14 y 19.1, mientras que en la población hispanoamericana fue el 19.1.



**Figura 23.** Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR D12S391.

### 5.3.20. Marcador D2S1338

El alelo de mayor frecuencia en la muestra poblacional peruana fue el alelo 19 (figura 24) con 0.225 y en la población hispanoamericana también fue este alelo con una frecuencia de 0.1928, por otro lado, el alelo de menor frecuencia alélica en la muestra poblacional peruana fue el 26, mientras que en la población hispanoamericana fue también el alelo 26.



**Figura 24.** Gráfico comparativo de las frecuencias alélicas de la muestra poblacional mestiza peruana y la hispanoamericana para el marcador STR D2S1338.



#### 5.4. Análisis de parámetros forenses

Los marcadores analizados (tabla 4) mostraron un valor de Índice de Contenido Polimórfico (PIC) mayor a 0.5, con un rango de 0.5315 (D22S1045) a 0.8534 (FGA), siendo los de más alto valor D18S51 (0.8318), D1S1656 (0.8400), FGA (0.8534) y los de menor valor D22S1045 (0.5315), TPOX (0.5406), D2S441 (0.5603). En cuanto al Poder de Discriminación (PD) este se encontró en el rango de 0.7620 (D22S1045) a 0.9665 (FGA), siendo los de más alto valor D2S13338 (0.9586), D1S1656 (0.9597), FGA (0.9665) y los de más bajo valor D22S1045 (0.7620), TPOX (0.7790).

A su vez el Poder de Exclusión se encontró en el rango de 0.2564 (TPOX) a 0.7047 (D1S1556), siendo los de más alto valor D19S433 (0.6559), D18S51 (0.6753), D1S1656 (0.7047) y los de más bajo valor fueron TPOX (0.2564), TH01 (0.3093) y D2S441 (0.3156) .

La Probabilidad de Coincidencia (PM) se encontró en el rango de 0.0336 (FGA) a 0.2381 (D22S1045), siendo los de más alto valor TPOX (0.2210), D22S1045 (0.2381) y los de más bajo valor FGA (0.0336), D1S1656 (0.0403), D2S1338 (0.0415).

El PD, PE y PM combinado (tabla 5) para estos 20 marcadores fue 0.999999999, 0.999999672, y  $2.532705772 \times 10^{-17}$ , respectivamente.

**Tabla 4.** Parámetros de interés forense.

	D3S1358	vWA	D16S539	CSF1PO	TPOX	D8S1179	D21S11	D18S51	D2S441	D19S433	TH01	FGA	D22S1045	D5S818	D13S317	D7S820	D10S1248	D1S1656	D12S391	D2S1338
<b>PD</b>	0.8162	0.8750	0.9156	0.8584	0.7790	0.9229	0.9559	0.9570	0.8106	0.9272	0.8550	0.9665	0.7620	0.8715	0.9233	0.8769	0.8669	0.9597	0.9345	0.9586
<b>PIC</b>	0.6006	0.6652	0.7558	0.6554	0.5406	0.7653	0.8267	0.8318	0.5603	0.7831	0.6423	0.8534	0.5315	0.6568	0.7634	0.6886	0.6549	0.8400	0.7847	0.8311
<b>PE</b>	0.3980	0.4283	0.6367	0.3980	0.2564	0.5625	0.5806	0.6753	0.3156	0.6559	0.3093	0.6462	0.3284	0.4439	0.5898	0.4928	0.4206	0.7047	0.6367	0.6462
<b>PM</b>	0.1839	0.1250	0.0844	0.1417	0.2210	0.0772	0.0442	0.0430	0.1895	0.0728	0.1451	0.0336	0.2381	0.1285	0.0768	0.1232	0.1332	0.0403	0.0655	0.0415

Siendo PD: Poder de Discriminación, PIC: Índice de Contenido Polimórfico, PE: Poder de Exclusión, PM: Probabilidad de Coincidencia.

**Tabla 5.** Parámetros de interés forense combinado.

	<b>PD</b>	<b>PE</b>	<b>PM</b>
<b>Combinado</b>	0.999999999	0.999999672	2.532705772e-17

Siendo PD: Poder de Discriminación, PE: Poder de Exclusión, PM: Probabilidad de Coincidencia.

## **5.5. Análisis de diferenciación poblacional**

El análisis de diferenciación poblacional se determinó comparando la muestra poblacional mestiza peruana, con otras poblaciones peruanas y mundiales. La diferenciación poblacional se realizó analizando primero 13 marcadores STRs y luego 20 marcadores STRs. Las poblaciones se encuentran rotuladas con números del 1 al 23 (para 13 STRs, tabla 6) y del 1 al 6 (para 20 STRs, tabla 8).

### **5.5.1. Análisis de diferenciación poblacional con 13 marcadores STRs entre 23 poblaciones**

Se analizaron los siguientes marcadores STRs: CSF1PO, FGA, TH01, TPOX, VWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51 y D21S11. La Tabla 6 presenta grupos poblacionales relacionados a la población mestiza peruana (1), como es la población de Perú, de Pérez *et al.*, 2003 (2) y la población amazónica de los Ashaninka (3); se analizó a su vez con poblaciones sudamericanas como Argentina (4), Bolivia (5), Brasil (6), Ecuador (7), Venezuela (8), Colombia- Bogotá (9), Colombia Boyacá (10), la población latina de México (11), grupos poblacionales de Afroamericanos (12), Asiáticos (13), Caucásicos (14) e Hispanoamericanos (15 y 16) de pobladores residentes en los Estados Unidos (EE.UU.). Adicionalmente se comparó con poblaciones caucásicas europeas como España (17), Rusia (18), Croacia (19), con poblaciones Asiáticas de China (20) y Japón (21) y poblaciones Africanas de Guinea (22) y Uganda (23).

Para la determinación de subdivisión poblacional entre las poblaciones analizadas, se tomó como valor de referencia  $F_{st}=0.01$ , el cual es un umbral conservativo para la mayoría de grupos poblacionales (Budowle *et*

*al.*, 2001; De la cruz, 2006). De tal manera que valores por debajo de este umbral indican que no hay subdivisión poblacional, y valores por encima de este indican la presencia de subdivisión o estructuración poblacional.

En el análisis de los 13 marcadores STRs, los valores de  $F_{st}$  promedio obtenidos entre la muestra poblacional mestiza peruana y las otras poblaciones peruanas y mundiales (tabla 7), indican que no existe subdivisión genética entre la muestra poblacional mestiza peruana y los grupos poblacionales de Perú (2) ( $F_{st}=0.003765$ ), de Bolivia (5) ( $F_{st}=0.005085$ ), y de Ecuador (7) ( $F_{st}=0.008846$ ) ya que los valores de  $F_{st}$  están por debajo del umbral conservativo 0.01, por otro lado la comparación con las muestras poblacionales de hispanoamericanos residentes en los Estados Unidos (15) (16) ( $F_{st}= 0.017670$ ,  $F_{st}=0.022233$ , respectivamente) determino que existe subdivisión genética entre ambas poblaciones, ya que los valores de  $F_{st}$  fueron mayores a 0.01, además se observó subdivisión genética con la población Ashaninka del Perú (3) ( $F_{st}=0.033649$ ), las poblaciones sudamericanas (4, 6, 8, 9, 10) ( $F_{st}=0.018707$ ,  $F_{st}=0.028602$ ,  $F_{st}=0.048125$ ,  $F_{st}=0.013937$ ,  $F_{st}=0.017393$ , respectivamente), la población de México (11) ( $F_{st}=0.011391$ ), las poblaciones residentes de los EE.UU. (12, 13, 14, 15, 16) ( $F_{st}=0.046858$ ,  $F_{st}=0.038454$ ,  $F_{st}=0.035728$ ,  $F_{st}=0.017670$ ,  $F_{st}=0.022233$ , respectivamente), las poblaciones caucásicas de Europa (17, 18, 19) ( $F_{st}=0.036735$ ,  $F_{st}=0.038420$ ,  $F_{st}=0.038879$ , respectivamente), las poblaciones Asiáticas (20, 21) ( $F_{st}=0.039657$ ,  $F_{st}=0.034382$ , respectivamente), y las poblaciones africanas (22, 23) ( $F_{st}=0.059898$ ,  $F_{st}=0.057762$ , respectivamente). En la figura 25 se observa de una manera más esquemática las diferencias en los valores

Fst para las 23 poblaciones, teniendo una coloración blanca aquellas poblaciones cuyo valor de Fst es cercano a cero, una coloración celeste cuando se aproxima a 0.01 y una coloración negra cuando su valor es 0.1.

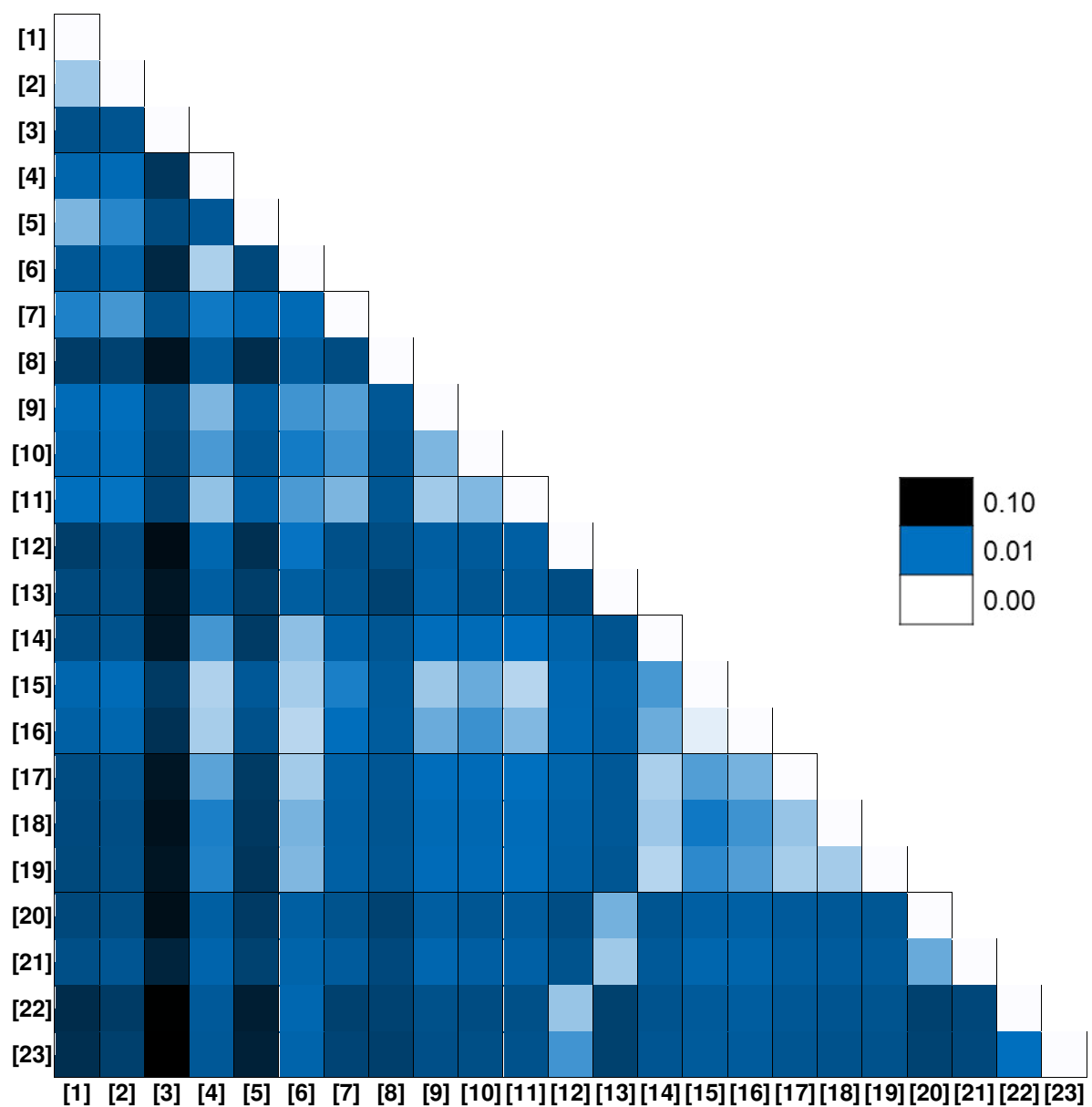
La figura 26 representa el árbol UPGMA basado en las distancias genéticas de Reynolds obtenida de los 13 marcadores STRs. Como se observa en el árbol, la muestra poblacional mestiza peruana se agrupa con la población peruana de Pérez *et al.* (2003) y la población de Bolivia, mientras la población Asháninka queda aislada de este grupo. De otro lado, las poblaciones hispanoamericanas reportadas por Butler *et al.* (2003) y Hill *et al.* (2013) tienden a agruparse entre ellas y con las poblaciones de Brasil, Argentina, Colombia-Bogotá, Colombia- Boyacá, México y Ecuador. En cambio, la población de Venezuela tiende a aproximarse más a las poblaciones africanas; asimismo se aprecia como las poblaciones caucásicas, asiáticas y africanas tienden a agruparse de acuerdo a su origen étnico.

**Tabla 6.** Poblaciones utilizadas para el análisis de 13 marcadores STRs, enumeradas del 1 al 23.

ETIQUETA	POBLACIONES
1	Mestizos peruanos (n=200, 21 departamentos y el Callao)
2	Perú (Pérez <i>et al.</i> , 2003) (n=100, Lima, Huancayo, Iquitos, Piura, Huaraz)
3	Asháninka (Tineo, 2017) (n= 181, Oxapampa-Pasco)
4	Argentina (Bozzo <i>et al.</i> , 2007) (n=154-203, Buenos Aires)
5	Bolivia (Rocabado <i>et al.</i> , 2009) (n=200, La Paz, mestizos)
6	Brasil (Aguiar <i>et al.</i> , 2014) (n=93 350-96 395, todos los estados)
7	Ecuador (González-Andrade <i>et al.</i> , 2003) (n=317, mestizos)
8	Venezuela (Chiurillo <i>et al.</i> , 2003) (n=255, Caracas)
9	Colombia- Bogotá (Rey <i>et al.</i> , 2003) (n=113-153)
10	Colombia- Boyacá (Rey <i>et al.</i> , 2003) (n=123-155)
11	México (Martínez-Sevilla <i>et al.</i> , 2016) (n=374, occidente de México)
12	Afroamericanos EE.UU. (Hill <i>et al.</i> , 2013) (n=342)
13	Asiáticos EE.UU. (Hill <i>et al.</i> , 2013) (n=97)
14	Caucásicos EE.UU. (Hill <i>et al.</i> , 2013) (n=361)
15	Hispanoamericanos EE.UU. (Hill <i>et al.</i> , 2013) (n=236)
16	Hispanoamericanos EE.UU. (Butler <i>et al.</i> , 2003) (n=140)
17	España (Camacho <i>et al.</i> , 2007) (n=329-342, caucásicos)
18	Rusia (Semikhodskii <i>et al.</i> , 2012) (n=1098-1118, caucásicos)
19	Croacia (Projic <i>et al.</i> , 2007) (n=195, caucásicos)
20	China (Liu <i>et al.</i> , 2017) (n=214, Shaanxi del sur)
21	Japón (Tie <i>et al.</i> , 2006) (n=650, Tokio)
22	Guinea (Alves <i>et al.</i> , 2005) (n=134)
23	Uganda (Gomes <i>et al.</i> , 2009) (n=218, Karamoja)

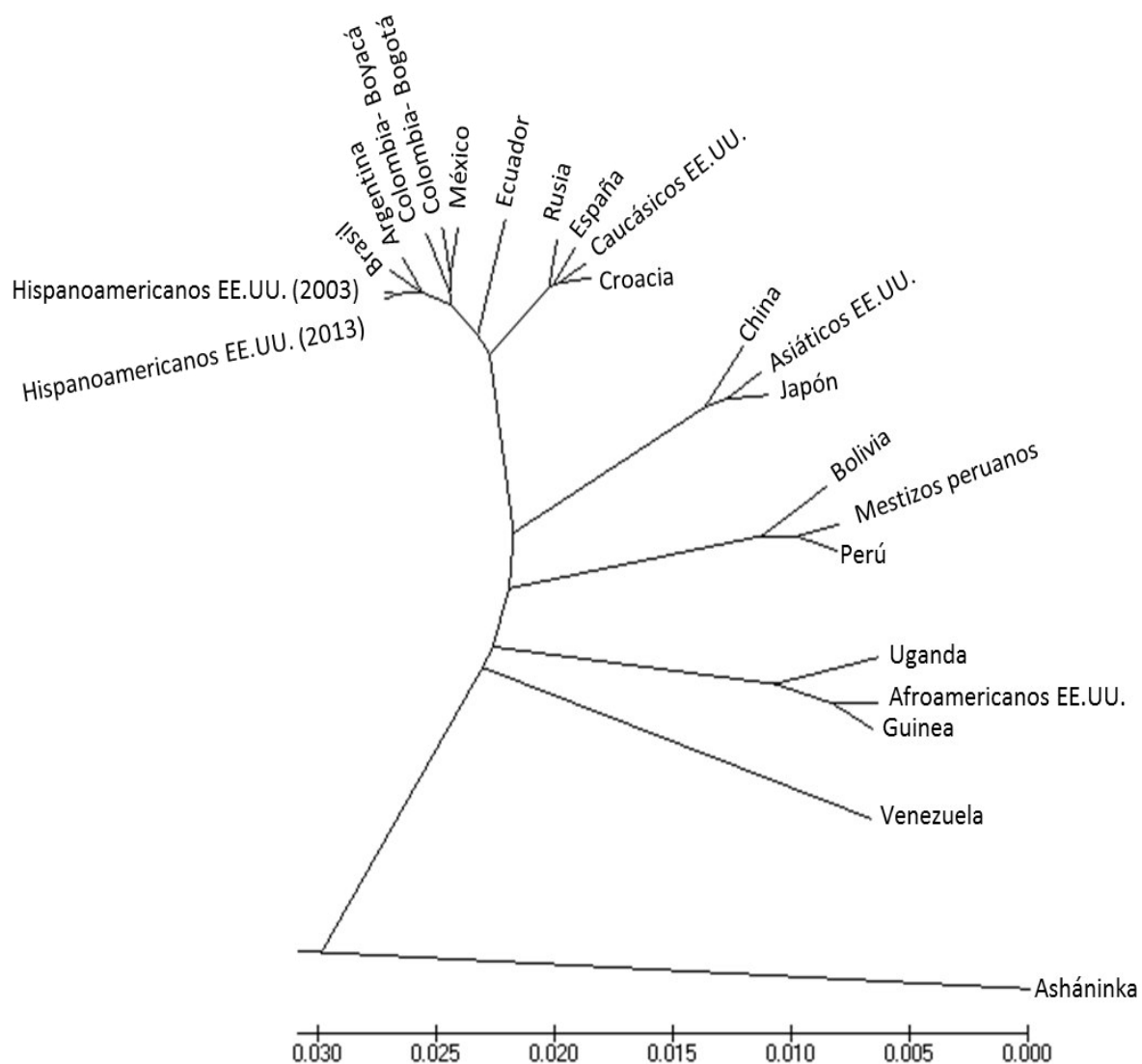
**Tabla 7.** Valores Fst promedio obtenidos en el análisis entre pares de poblaciones para los 13 marcadores STRs. Los resultados resaltados en letra negrita muestran los valores Fst entre la muestra poblacional mestiza peruana (1) y el resto de poblaciones analizadas.

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]	[17]	[18]	[19]	[20]	[21]	[22]	[23]
[1]	0.000000																						
[2]	0.003765	0.00000																					
[3]	0.033649	0.03080	0.00000																				
[4]	0.018707	0.01509	0.05284	0.00000																			
[5]	0.005085	0.00846	0.03747	0.02824	0.00000																		
[6]	0.028602	0.02301	0.06291	0.00320	0.03963	0.00000																	
[7]	0.008846	0.00732	0.03319	0.00942	0.01666	0.01499	0.00000																
[8]	0.048125	0.04377	0.07771	0.02595	0.05923	0.02496	0.03645	0.00000															
[9]	0.013937	0.01197	0.04026	0.00502	0.02402	0.00748	0.00677	0.02862	0.00000														
[10]	0.017393	0.01415	0.04331	0.00707	0.02858	0.00925	0.00754	0.03077	0.00504	0.00000													
[11]	0.011391	0.00984	0.04280	0.00421	0.02104	0.00706	0.00511	0.02963	0.00362	0.00488	0.00000												
[12]	0.046858	0.03710	0.08286	0.01724	0.05678	0.00980	0.03350	0.03614	0.02317	0.02644	0.02271	0.00000											
[13]	0.038454	0.03564	0.07591	0.02358	0.04657	0.02363	0.03084	0.04378	0.02131	0.03007	0.02643	0.03567	0.00000										
[14]	0.035728	0.03211	0.07526	0.00734	0.04864	0.00440	0.02032	0.02950	0.01221	0.01388	0.01082	0.02056	0.03050	0.00000									
[15]	0.017670	0.01416	0.04977	0.00308	0.02778	0.00346	0.00898	0.02590	0.00382	0.00585	0.00280	0.01676	0.02221	0.00719	0.00000								
[16]	0.022233	0.01800	0.05608	0.00341	0.03264	0.00272	0.01178	0.02524	0.00583	0.00769	0.00490	0.01597	0.02312	0.00576	0.00102	0.00000							
[17]	0.036735	0.03244	0.07594	0.00642	0.04921	0.00357	0.02127	0.02906	0.01225	0.01381	0.01058	0.01942	0.02767	0.00327	0.00675	0.00533	0.00000						
[18]	0.038420	0.03558	0.07951	0.00894	0.05092	0.00526	0.02252	0.03018	0.01474	0.01631	0.01305	0.02112	0.02778	0.00379	0.00944	0.00755	0.00403	0.00000					
[19]	0.038879	0.03530	0.07666	0.00873	0.05295	0.00493	0.02211	0.02951	0.01364	0.01518	0.01282	0.02208	0.02935	0.00284	0.00825	0.00679	0.00342	0.00350	0.00000				
[20]	0.039657	0.03616	0.08093	0.02309	0.04939	0.02305	0.03170	0.04354	0.02346	0.02944	0.02557	0.03577	0.00543	0.03005	0.02248	0.02213	0.02602	0.02780	0.02869	0.00000			
[21]	0.034382	0.02980	0.06592	0.01891	0.04413	0.01945	0.02534	0.03918	0.01731	0.02313	0.02186	0.03168	0.00371	0.02712	0.01794	0.01853	0.02434	0.02565	0.02614	0.00590	0.00000		
[22]	0.059898	0.04900	0.09120	0.02687	0.06997	0.01709	0.04472	0.04377	0.03306	0.03648	0.03390	0.00397	0.04481	0.03138	0.02623	0.02426	0.02841	0.03063	0.03083	0.04450	0.03987	0.00000	
[23]	0.057762	0.04564	0.09150	0.02780	0.06800	0.01926	0.04226	0.04582	0.03412	0.03467	0.03224	0.00744	0.04487	0.02984	0.02583	0.02562	0.03013	0.03232	0.03229	0.04327	0.03912	0.01102	0.00000



**Figura 25.** Heatmap para 13 marcadores STRs. Basado en el  $F_{st}$  promedio obtenido en el análisis entre pares de poblaciones para los 13 marcadores STRs de la tabla 7.





**Figura 26.** Árbol UPGMA para 13 marcadores STRs. Basado en las distancias de Reynolds a partir de la matriz de la tabla 7 para los 13 marcadores STRs y 23 poblaciones.

### **5.5.2. Análisis de diferenciación poblacional con 20 marcadores STRs entre 6 poblaciones**

Los 20 marcadores STRs analizados son: CSF1PO, FGA, TH01, TPOX, VWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, D1S1656, D2S441, D2S1338, D10S1248, D12S391, D19S433, D22S1045. La tabla 8 presenta la población de Brasil (2), los grupos poblacionales de Afroamericanos (3), Asiáticos (4), Caucásicos (5) e Hispanoamericanos (6) de pobladores residentes en los EE.UU.

Para la determinación de subdivisión poblacional entre las poblaciones analizadas, se tomó como valor de referencia  $F_{st}=0.01$ , el cual es un umbral conservativo para la mayoría de grupos poblacionales (Budowle *et al.*, 2001; De la cruz, 2006). De tal manera que valores por debajo de este umbral indican que no hay subdivisión poblacional, y valores por encima de éste indican la presencia de subdivisión o estructuración poblacional.

En el análisis para los 20 marcadores STRs los valores de  $F_{st}$  promedio obtenidos entre la muestra poblacional mestiza peruana y las poblaciones mundiales (tabla 9), indican que existe subdivisión poblacional entre la muestra poblacional mestiza peruana y la población hispanoamericana (6) ( $F_{st}=0.016713$ ), ya que los valores de  $F_{st}$  fueron mayores a 0.01, a su vez también se encontró subdivisión genética entre la población de Brasil (2) ( $F_{st}=0.029163$ ) y las poblaciones residentes de los EE.UU. (3, 4, 5) ( $F_{st}=0.051786$ ,  $F_{st}=0.038883$ ,  $F_{st}=0.035750$ , respectivamente). En la figura 27 se observa de una manera más perceptible las diferencias en los valores  $F_{st}$  para las 6 poblaciones, teniendo una coloración blanca aquellas

poblaciones cuyo valor de  $F_{st}$  es cercano a cero, una coloración celeste cuando se aproxima a 0.01 y una coloración negra cuando su valor es 0.1.

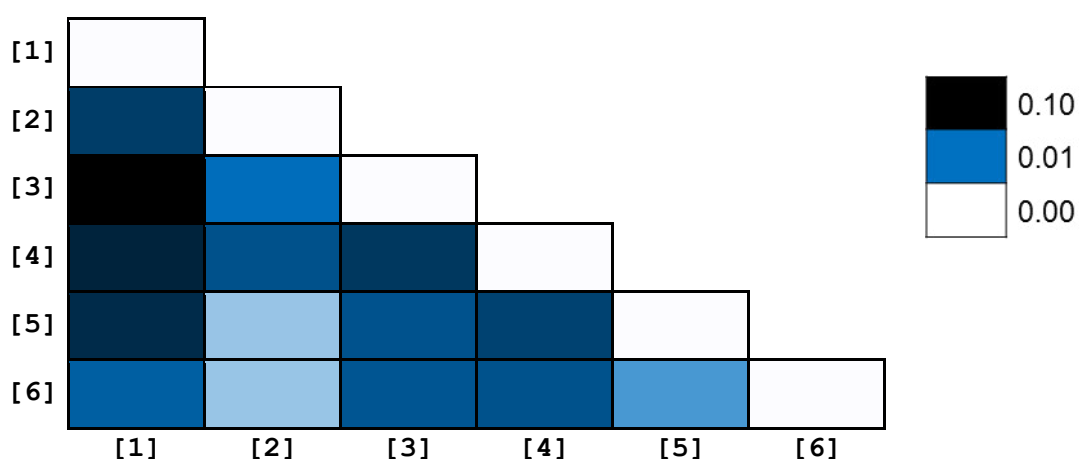
La figura 28 representa el árbol UPGMA basado en las distancias genéticas de Reynolds obtenida de los 20 marcadores STRs. Se observa que la muestra poblacional mestiza peruana se aísla de la población hispanoamericana, mientras que la población de Brasil tiende a agruparse con los caucásicos e hispanoamericanos, por otra parte las poblaciones afroamericanas y asiáticas tienden a estar separadas del resto de las poblaciones comparadas.

**Tabla 8.** Poblaciones, enumeradas del 1 al 6, utilizadas para el análisis de 20 marcadores STRs.

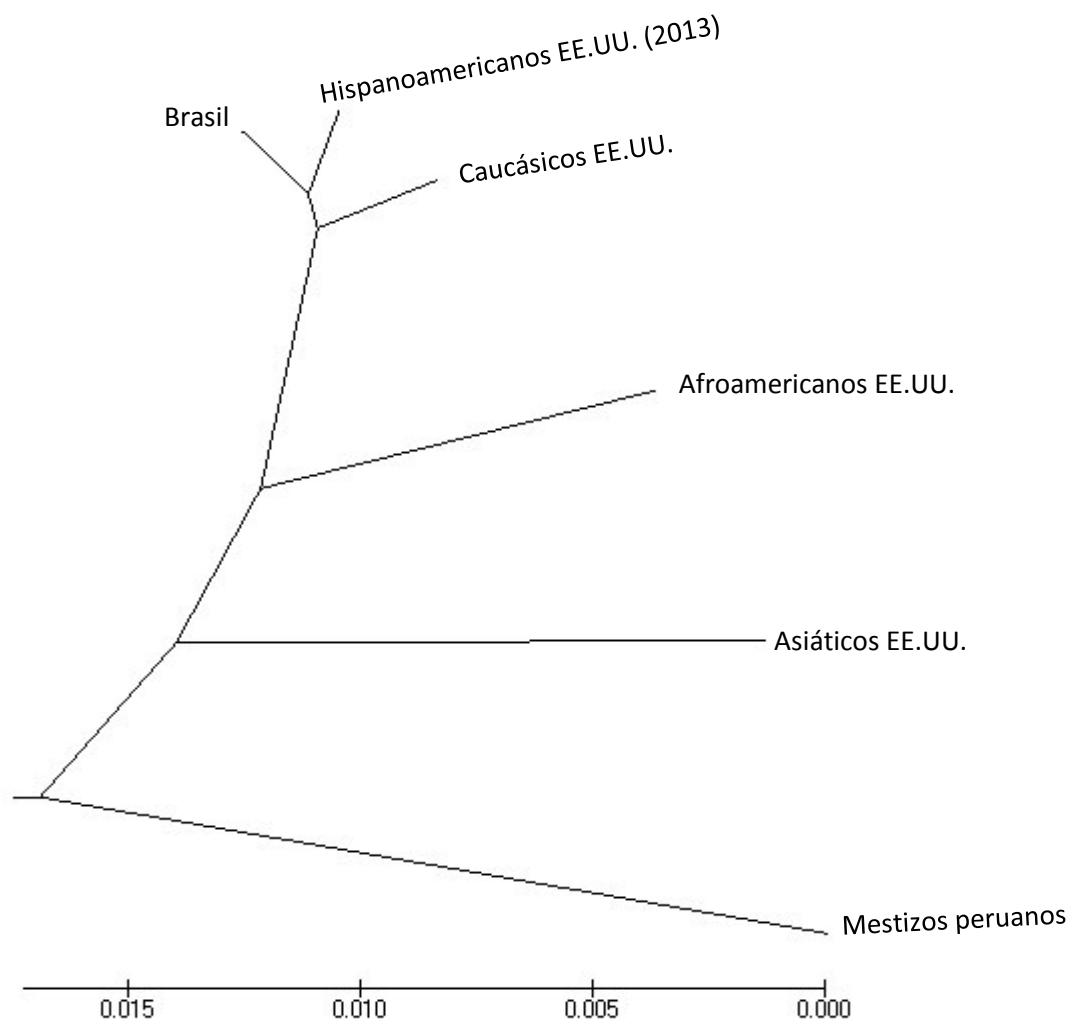
ETIQUETA	POBLACIONES
1	Mestizos peruanos (n=200, 21 departamentos y el Callao)
2	Brasil (Aguiar <i>et al.</i> , 2014) (n=93 350-96 395, todos los estados)
3	Afroamericanos EE.UU. (Hill <i>et al.</i> , 2013) (n=342)
4	Asiáticos EE.UU. (Hill <i>et al.</i> , 2013) (n=97)
5	Caucásicos EE.UU. (Hill <i>et al.</i> , 2013) (n=361)
6	Hispanoamericanos EE.UU. (Hill <i>et al.</i> , 2013) (n=236)

**Tabla 9.** Valores Fst promedio obtenidos en el análisis entre pares de poblaciones para los 20 marcadores STRs. Los resultados resaltados en letra negrita muestran los valores Fst entre la muestra poblacional mestiza peruana (1) y el resto de poblaciones analizadas.

	[1]	[2]	[3]	[4]	[5]	[6]
[1]	<b>0.000000</b>					
[2]	<b>0.029163</b>	0.000000				
[3]	<b>0.051786</b>	0.011152	0.000000			
[4]	<b>0.038883</b>	0.021686	0.031113	0.000000		
[5]	<b>0.035750</b>	0.004003	0.021252	0.027372	0.000000	
[6]	<b>0.016713</b>	0.003999	0.020199	0.021327	0.007157	0.000000



**Figura 27.** Heatmap para 20 marcadores STRs. Basado en el Fst promedio obtenido en el análisis entre pares de poblaciones para los 20 marcadores STRs de la tabla 9.



**Figura 28.** Árbol UPGMA para 20 marcadores STRs. Basado en las distancias de Reynolds a partir de la matriz de la tabla 9 para los 20 marcadores STRs y 6 poblaciones.

## 6. DISCUSIÓN

### 6.1. Parámetros poblacionales

#### 6.1.1. Tamaño y característica de la población

El Perú se caracteriza por ser una población mestiza como resultado de una mezcla principalmente entre nativos americanos (esencialmente Quechuas y Aymaras), españoles y en menor grado africanos y asiáticos (Talledo *et al.*, 2010). Homburger *et al.* (2015) detallan en su estudio genético realizado en la población peruana que ésta contiene un componente ancestral de 68.3% de nativo americano, 26% de europeo, 3.2% de africano, y 2.5% de asiático. Además la *National Geographic* analizó muestras recolectadas de peruanos que vivían en la ciudad de Lima. En el mencionado trabajo destacan que las personas en esta región son predominantemente nativas americanas (68%) con una significativa cantidad de la más reciente mezcla con poblaciones europeas en los últimos 500 años durante la era colonial española (comprendiendo un componente del 15% del Mediterráneo, del 10% del norte de Europa y 3% del sudoeste asiático). También se expresa que hay una pequeña cantidad de África subsahariana (2%), debido a la importante trata de esclavos africanos en los siglos XVI-XIX. Por último el censo realizado por el Instituto Nacional de Estadística e Informática (INEI) en el 2017, indicó que el 60.2% de la población peruana se autoidentifica como mestizo, esto en relación a cómo se consideran o autoperceben de acuerdo a sus costumbres y antepasados.

En el presente trabajo se realizó un análisis de una muestra poblacional de 200 individuos mestizos no emparentados residentes en la ciudad de Lima y/o Callao, el cual como se menciona líneas arriba se caracteriza por ser descendientes de nativos americanos, españoles y en menor medida de africanos y asiáticos.

Existen diferentes opiniones de los especialistas en relación al número de individuos que deben ser muestreados en una población a fin de que ésta sea representativa. Evett y Gill (1991) enfatizan que de 100 a 120 individuos por locus por población son suficientes para realizar cálculos robustos de probabilidad. Chakraborty (1992) señala que entre 100 a 150 individuos por población puede proporcionar un adecuado muestreo para un locus genético siempre que las frecuencias alélicas por debajo del 1% no sean usadas en los cálculos forenses. Otros autores sostienen que para que los resultados sean significativos el tamaño de la población debe ser como mínimo de 200 individuos (Fernández, 2008). Por lo tanto, cuanto más grande sea la base de datos más representativa de la población ésta será. La práctica actual dicta que varios cientos de individuos deberían ser muestreados cuando se crea una base de datos de frecuencias alélicas.

Otra característica importante es que estas personas muestreadas no deberían estar relacionadas directamente; por consiguiente, no ser hermanos o madres y niños, de tal manera que las combinaciones de estas no deben ser incorporadas en una base de datos de frecuencias alélicas. (Goodwin *et al.*, 2011).

La población analizada en este trabajo cumple con las características mencionadas, ninguno de los 200 individuos están emparentados. Asimismo, es muy importante resaltar que en este muestreo, se encontró individuos que se les podría considerar representantes de varios departamentos, ya que la población analizada estuvo conformada por pobladores de los 20 departamentos del país más la provincia constitucional del Callao (Anexo 3); sin embargo, en algunos voluntarios de la población de Lima y Callao (18), los padres y abuelos paternos y maternos son de un mismo departamento distinto al lugar de nacimiento del voluntario (Lima), muy probablemente debido a que los padres emigraron a la capital por razones de trabajo y/o búsqueda de mejores oportunidades. Por esta razón se les considero a los 18 voluntarios nacidos en Lima y/o Callao (Anexo 4), como lugar de origen el departamento en el cual todos sus ancestros nacieron. Así se tiene por ejemplo que los voluntarios 29 y 55 nacieron en Lima pero las generaciones anteriores son de Cusco y de Huancavelica, respectivamente (Anexo 5). Es así que realmente la población analizada en este trabajo, abarco un total de 21 departamentos más la provincia constitucional del Callao. En ese sentido, este muestreo es una representación no solo de Lima sino de varios departamentos del Perú.



### 6.1.2. Análisis de los parámetros poblacionales

En el presente estudio de la muestra poblacional mestiza peruana se han encontrado alelos no reportados en trabajos anteriores como la población de Pérez *et al.* (2003) y Barletta (2010), y también alelos reportados por esos autores pero no encontrados en la muestra poblacional mestiza peruana del presente estudio. Es así que se resalta lo siguiente:

Para el marcador D3S1358 los alelos 11, 12 y 13 no fueron hallados en la población de Pérez *et al.* (2003), mientras que en la población de Barletta (2010) no se encontró el alelo 12, en cambio sí han sido hallados en la muestra poblacional mestiza peruana trabajada.

En relación al marcador vWA, no se registró el alelo 13 en la población de Pérez *et al.* (2003) ni en el trabajo de Barletta (2010) pero en la población de Pérez *et al.* (2003) se encontró el alelo 12, en cambio Barletta (2010) reporta el alelo 21; ambos alelos no han sido encontrados en este trabajo.

Considerando el marcador D16S539, la población de Pérez *et al.* (2003) no registro el alelo 15, el cual si fue hallado en la muestra poblacional mestiza peruana.

Analizando el marcador CSF1PO los alelos 7, 8, 12.3 y 15 no fueron hallados en la población de Pérez *et al.* (2003), mientras que la población de Barletta (2010) no registro el alelo 12.3, sin embargo estos alelos si han sido determinados en el presente trabajo.

Examinando el marcador TPOX, los alelos 6 y 8.2 no fueron encontrados en la población de Pérez *et al.* (2003), mientras que la población de

Barletta (2010) no registro el alelo 8.2, estos alelos fueron determinados en este trabajo; en cambio, en la población de Barletta (2010) se encontró el alelo 7, que no ha sido registrado en este estudio.

Al observar el marcador D8S1179 el alelo17 no fue hallado en la población de Pérez *et al.* (2003), en cambio este alelo si fue hallado en la muestra poblacional mestiza peruana, por otro lado Pérez *et al.* (2003) y Barletta (2010) reportaron el alelo 8 que no fue encontrado en la muestra poblacional mestiza peruana.

En relación al marcador D21S11 los alelos 29.2 y 35 no fueron hallados en la población de Pérez *et al.* (2003), sin embargo Pérez registro el alelo 26 que no fue encontrado en la muestra poblacional mestiza peruana.

Al analizar en relación al marcador D18S51, los alelos 22 y 25 no fueron hallados en la población de Pérez *et al.* (2003), mientras que en la población de Barletta (2010) no se encontró el alelo 25, sin embargo en el trabajo de Barletta se encontró los alelos 23 y 24 que no han sido encontrados en la muestra poblacional mestiza peruana de esta investigación.

El marcador TH01 en las poblaciones de Pérez *et al.* (2003) y Barletta (2010) se encontró el alelo 10 el mismo que no ha sido registrado en este estudio; para el marcador FGA en la población de Pérez *et al.* (2003) no se encontró el alelo 22.2, sin embargo sí registro el alelo 17 que no fue encontrado en la muestra poblacional mestiza peruana analizada en este trabajo.

Con respecto a la población de Pérez *et al.* (2003) se observa que para el marcador D5S818, no se encontró el alelo 8; para el marcador D13S317 el alelo 9.2 y 17 no fueron encontrados y para el marcador D7S820 no se encontró el alelo 14; sin embargo, estos alelos sí fueron encontrados en la muestra poblacional mestiza peruana del presente trabajo.

Esta variabilidad con respecto al trabajo de Pérez *et al.* (2003) quizá se deba al tamaño poblacional trabajado, ya que en dicho estudio se analizó una población de 100 individuos, abarcando las ciudades de Huancayo, Iquitos, Piura, Huaraz y Lima, mientras que en el presente estudio se analizó 200 individuos residentes en Lima y/o Callao, pero como se menciona líneas arriba son provenientes de los 20 departamentos del país más la provincia constitucional del Callao (Anexo 3) teniendo por lo tanto un acervo genético más amplio. Igualmente, las diferencias encontradas con el trabajo de Barletta (2010) es muy probable que también se deba a la procedencia de su población, aunque este dato no está detallado en el reporte.

Se resalta que en el presente trabajo se detectaron 8 alelos que todavía no han sido descritos en la base de datos del *THE NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY* (NIST), el cual es una base de datos internacional que recopila y mantiene información acerca de los marcadores STRs para la caracterización forense (Ruitberg *et al.*, 2001). Estos alelos fueron 12.3 del marcador CSF1PO, 8.2 del marcador TPOX, 9.1 del marcador D2S441, 11.2 del marcador D19S433, 9.2 del marcador D13S317 y 14, 19.1, 20.3 del marcador D12S391.

En el análisis de diversidad poblacional Butler (2005) señala que una alta heterocigosidad significa que existe una mayor diversidad de alelos y, por lo tanto, hay menos posibilidades de una coincidencia aleatoria en la muestra; para que los marcadores STRs se consideren para fines de identificación humana deben de tener una Het (obs) > 70%. Según los resultados de este trabajo, los marcadores como vWA (0.700), D16S539 (0.8200), D8S1179 (0.7800), D21S11 (0.7900), D18S51 (0.8400), D19S433 (0.8300), FGA (0.8250), D5S818 (0.7100), D13S317 (0.7950), D7S820 (0.7400), D1S1656 (0.8550), D12S391 (0.8200), D2S1338 (0.8250) presentaron valores de Het (obs) mayores a 0.7, razón por lo cual son considerados óptimos para identificación humana.

De otro lado, la Heterocigosidad esperada o diversidad génica, permite dar un cálculo más exacto de la heterocigosidad en la población, siendo un excelente estimador de la variabilidad, asimismo, valores inferiores a 0.5 para la Het (esp) se consideran insuficientes (Filho, 2012). Todos los locis analizados en este trabajo muestran valores de heterocigosidad esperada mayor a 0.5900, teniendo un valor promedio de 0.7489, considerándose que los 20 STRs analizados presentan una gran diversidad génica.

Cuando la Heterocigosidad observada es igual a la heterocigosidad esperada se dice que la población se encuentra en equilibrio de Hardy-Weinberg (Gattepaille, 2015). De los 20 marcadores trabajados, 19 resultaron en equilibrio de H-W. Solo el locus D19S433 ( $p=0.0382$ ), no se encontró en equilibrio razón por la cual se aplicó la corrección de Bonferroni. Esta fórmula permite ajustar el nivel de rechazo para el más pequeño valor de *P-value* de un nivel general de  $\alpha=0.05$  a  $0.05/x$ , donde  $x$

es igual al número de loci (Hameed *et al.*, 2015). La corrección de Bonferroni es usada para neutralizar las múltiples comparaciones en la prueba del Equilibrio de Hardy-Weinberg (Pereira *et al.*, 2018). Se obtuvo como nuevo valor de nivel de significancia 0.0025, con lo cual desaparece las desviaciones encontradas y considerándose a todos estos marcadores en equilibrio de H-W. El hecho de que los locis estén en equilibrio de Hardy-Weinberg permite el uso de la fórmula cuadrática binomial, que es usada para estimar la frecuencia genotípica (Martínez-Sevilla *et al.*, 2016; Orellana *et al.*, 2017).

La prueba de desequilibrio de ligamiento (DL) es necesario para descartar asociación entre pares de loci, pudiendo indicar independencia de los loci, haciendo posible el uso del producto de la regla para estimar la frecuencia de los perfiles genéticos o para calcular los índices de paternidad (Martínez-Sevilla *et al.*, 2016; Orellana *et al.*, 2017). El desequilibrio de ligamiento fue usado para todas las posibles combinaciones entre los 20 locis. Once pares de loci (D3S1358- D10S1248, vWA- D2S1338, D16S539- D7S820, CSF1PO- TH01, TPOX- D19S433, D21S11- D13S317, D21S11- D7S820, D21S11- D12S391, D18S51- TH01, D18S51- D1S1656, D13S317- D2S1338) presentaron desequilibrio de ligamiento ( $p < 0.05$ ) entre 190 pares de combinaciones. Después de aplicar la corrección de Bonferroni ningún par de loci exhibió DL (a un nivel de significancia = 0.000263,  $\alpha = 0.05/190$ ), además estos marcadores no se encuentran localizados en un mismo cromosoma por lo que han sido considerados genéticamente no ligados. Por lo tanto los 20 marcadores pueden tratarse como loci independientes a nivel de la población.

La frecuencia mínima determinada para este estudio fue de 0.0125, las frecuencias mínimas de los alelos se utilizan para evitar probabilidades alélicas muy pequeñas cercanas a cero, que pueden atribuirse al tamaño de la muestra (Restrepo *et al.*, 2011) ya que el uso de una frecuencia mínima permite la compensación de un muestreo poco disperso de alelos pocos frecuentes en una base de datos poblacional (Budowle *et al.*, 1996). La frecuencia alélica mínima es esencial para interpretar casos de alelos raros o alelos nulos que pueden ser usados como beneficio para el acusado (Orellana *et al.*, 2017), permitiendo al acusado obtener el beneficio de la duda (Martínez-Sevilla *et al.*, 2016).

En comparación con otras investigaciones realizadas en poblaciones peruanas, el trabajo de Pérez *et al.* (2003) realizó un estudio de 13 marcadores STRs (CSF1PO, FGA, TH01, TPOX, VWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51 y D21S11) del CODIS en pobladores del Perú de las ciudades de Huancayo, Iquitos, Piura y Huaraz, encontrando que todos los marcadores se encontraban en equilibrio de H-W. De los 13 marcadores STRs analizados solo 9 tenían altos valores de Het (obs) (mayores a 0.7), los locis TPOX, CSF1PO y vWA presentaron bajos valores de Het (obs), siendo los marcadores TPOX y CSF1PO los que coincidieron con este estudio.

Barletta (2010) investigó 7 marcadores STRs (CSF1PO, TPOX, TH01, vWA, D18S51, D3S1358, D8S1179) conformada por una muestra poblacional de la ciudad de Lima, determinando que todos los marcadores se encontraban en equilibrio de H-W, y que no presentaban DL. Igualmente, la Het (obs) para 5 de los 7 marcadores fue alta (mayores a 0.7) siendo los loci TPOX y TH01 los que presentaron baja Het (obs),

estos resultados están en correlación con el presente estudio. A su vez los resultados de Het (esp) mostraron valores altos (mayores a 0.5) coincidiendo también con los resultados obtenidos en este estudio.

Al realizar la comparación de alelos y frecuencias alélicas para la muestra poblacional mestiza peruana analizada en esta investigación y la población hispanoamericana de Hill *et al.* (2013), se encontró que la muestra en este estudio posee alelos que no registra la población hispanoamericana, es decir la muestra estudiada presenta alelos propios, siendo estos alelos los siguientes: los alelos 11 y 12 para el marcador D3S1358, el alelo 13 para el marcador vWA, el alelo 12.3 para el marcador CSF1PO, los alelos 8.2 y 13 para el marcador TPOX, los alelos 34.2 y 35 para el marcador D21S11, el alelo 25 para el marcador D18S51, el alelo 9.1 para el marcador D2S441, los alelos 11.2, 14.1 y 17 para el marcador D19S433, el alelo 19 para el marcador D22S1045, los alelos 9.2 y 17 para el marcador D13S317, el alelo 14 para el marcador D7S820, el alelo 9 para el marcador D10S1248, los alelos 14 y 20.3 para el marcador D12S391. A su vez la población hispanoamericana posee alelos propios que la caracterizan, siendo estos: el alelo 20 para el marcador D3S1358, los alelos 11 y 21 para el marcador vWA, el alelo 7 para el marcador TPOX, los alelos 8 y 18 para el marcador D8S1179, los alelos 24.2, 26.2, 33, 33.1, 34 para el marcador D21S11, los alelos 14.2 y 24 para el marcador D18S51, los alelos 12.3 y 17 para el marcador D2S441, los alelos 9 y 10 para el marcador D19S433, el alelo 10 para el marcador TH01, los alelos 17, 23.2, 29, 30 para el marcador FGA, los alelos 10 y 13 para el marcador D22S1045, el alelo 15 para el marcador D5S818, el alelo 10.3 para el marcador D7S820, el alelo 19 para el marcador D10S1248, el alelo 14.3 para el marcador D1S1656, los alelos 26 y 27

para el marcador D12S391, por consiguiente también se encontró variación en las frecuencias alélicas, ya que ciertos alelos en la muestra poblacional mestiza peruana tuvieron mayor o menor valor de frecuencia alélica en relación con la población hispanoamericana.

## **6.2. Parámetros forenses**

El índice de contenido polimórfico (PIC), poder de discriminación (PD), y poder de exclusión (PE) son utilizados para conferir mayor claridad en los análisis de vínculo genético y estudios poblacionales (Filho, 2012). Por otro lado, en el análisis de los parámetros forenses Butler (2005) sostiene que para que los marcadores STRs sean utilizados en la identificación humana estos deben de tener un  $PD > 90$ . En esta investigación los locis D16S539 (0.9156), D8S1179 (0.9229), D21S11 (0.9559), D18S51 (0.9570), D19S433 (0.9272), FGA (0.9665), D13S317 (0.9233), D1S1656 (0.9597), D12S391 (0.9345), D2S1338 (0.9586) presentaron un PD mayor a 0.9, considerándose como marcadores con un alto poder de discriminación. Al comparar nuestros resultados con el estudio de Pérez *et al.* (2003) se encontró que los marcadores con un alto PD fueron (mayor al 90%) FGA, D13S317, D8S1179, D21S11, D18S51, D16S539, teniendo un comportamiento similar a la población investigada en este estudio, además en el trabajo realizado por Barletta (2010), se determinó que los locis D18S51 y D8S1179 presentan un alto PD, entrando en concordancia con los resultados obtenidos en este trabajo.

Teniendo en cuenta que el PD es proporcional a las frecuencias alélicas y que los marcadores con mayor índice de PD son aquellos que presentan alelos con baja frecuencia (Filho, 2012), los resultados obtenidos en este trabajo indicarían que los marcadores que presentaron un  $PD < 90\%$  se encuentran poco distribuido en cuanto a sus frecuencias alélicas, ya que en la mayoría de



los marcadores STRs la suma de sus frecuencias alélicas de solamente 2 alelos daba como resultado una frecuencia mayor al 70 %, viéndose que los alelos restantes presentaban frecuencias alélicas muy bajas en relación a estos dos alelos. Por otra parte el PD combinado fue de 0.999999999, estableciéndose que los 20 marcadores STRs trabajando en conjunto son buenos discriminantes.

En cuanto al PIC Boltstein *et al.* (1980), plantea que marcadores con valores superiores a 0.5 son marcadores altamente informativos, valores de 0.25 a 0.5 medianamente informativos, y valores menores a 0.25 poco informativos, en este estudio los 20 locis analizados presentaron un PIC mayor a 0.5, mostrando un buen contenido de información polimórfica. Estos resultados concuerdan con el trabajo de Barletta (2010) ya que sus 7 marcadores STRs (CSF1PO, TPOX, TH01, vWA, D18S51, D3S1358, D8S1179) analizados obtuvieron altos valores de PIC.

Se realizó el cálculo del PEcum obteniéndose un valor 0.999999672 (99.9999672%), siendo la probabilidad de exclusión deseable la máxima posible. El Grupo Español y Portugués de la Sociedad Internacional de Genética Forense (GEP-ISFG) considera que para que un Laboratorio se considere competente en la realización de pruebas de paternidad, el PE debe de estar por encima del 99.9% (Fernández, 2008), el cual se ha conseguido en este estudio.

### 6.3. Análisis de diferenciación poblacional

El Perú es un país multiétnico, multilingüe, y pluricultural (Tineo, 2017). La población peruana ha tenido un gran acervo genético a lo largo de su historia comenzando con los primeros europeos que llegaron a Perú en el siglo XVI que fueron principalmente de España, quienes también trajeron a algunos africanos como esclavos. En 1849 comenzó una inmigración de China a todas las regiones del Perú para trabajar en plantaciones y explotación de guano, y desde 1899 también había algunos inmigrantes japoneses (Sandoval *et al.*, 2013). En 1853, algunas familias alemanas emigraron con el objetivo de colonizar la región amazónica, pero un gran número de europeos procedentes de Italia y otros países llegaron a principios del siglo XX, en particular durante la primera guerra mundial y principios de la segunda (1918–1938). Durante la primera década del siglo XX, se produjo un importante flujo de migración interna en la Amazonía peruana, cuando muchas comunidades urbanas e indígenas fueron desplazadas de su tierra de origen para obtener ganancias o escaparse del auge de la industria del caucho; sin embargo, desde 1940 se produjo un gran movimiento migratorio dentro del Perú, principalmente a Lima procedente de Junín, Ayacucho, La Libertad, Ica, Lambayeque, Cajamarca, Piura y, en menor grado, de otros lugares. Esta migración interna de fines del siglo XX estuvo compuesta principalmente por personas rurales e indígenas que se mudaron a ciudades urbanizadas, por lo que se esperaría un gran impacto en la ascendencia genómica de los habitantes de grandes centros urbanos como Lima (Sandoval *et al.*, 2013). Como se ve la composición étnica del Perú está influenciada por varios factores en los que se destaca la inmigración; en donde los grupos étnicos presentes en el Perú están constituidas por mestizos, amerindios, blancos, y en menor medida afro descendientes y asiáticos (Tineo, 2017).

A nivel interpoblacional la validación de los STRs incluye la comparación con otras poblaciones para establecer sus relaciones genéticas, estructura e incluso el conocimiento de sus orígenes (Orellana *et al.*, 2017). Para este trabajo se utilizó el estadístico  $F_{st}$  el cual es una de las más ampliamente usadas mediciones para diferenciación genética y desempeña un rol central en estudios genéticos (Liu *et al.*, 2017), ya que mide la variación de las frecuencias alélicas entre poblaciones, y por tanto la diferenciación genética entre ellas (Eguiarte *et al.*, 2007).

Los resultados de  $F_{st}$  obtenidos para el análisis de 13 marcadores STRs refleja la no existencia de diferenciación poblacional entre la muestra mestiza peruana de este estudio y el trabajo de Pérez *et al.* (2003), el cual abarcó pobladores de las ciudades de Huancayo, Iquitos, Piura y Huaraz, mientras que el presente estudio recopiló muestras de individuos de distintas partes del país que residían en la ciudad de Lima (Anexo 4), por lo que se podría considerar que ambas poblaciones presentan un mismo origen étnico. En cambio al comparar el estadístico  $F_{st}$  con la población hispanoamericana de Butler *et al.* (2003) y de Hill *et al.* (2013), los resultados obtenidos fueron diferentes, evidenciándose subdivisión poblacional entre la muestra poblacional mestiza peruana y la hispanoamericana, esto se puede corroborar aún más por medio de las gráficas de *Heatmap* (figura 24) donde se visualiza con una coloración más clara (celeste claro) las poblaciones que no presentan subdivisión poblacional, mientras que tonalidades más oscuras (azul) indican que son poblaciones que presentan diferenciación genética como es el caso de las poblaciones hispanoamericanas. De otro lado poblaciones como las africanas presentan una coloración oscura evidenciando una alta diferenciación. Resultados similares fueron obtenidos al hacer el análisis  $F_{st}$  para los 20 marcadores STRs (figura 25), evidenciándose claramente una

estructuración poblacional entre la muestra poblacional mestiza peruana y la población hispanoamericana.

Asimismo en el presente estudio se eligió edificar el árbol filogenético en función de las distancias de Reynolds, ya que estas permiten según De la cruz (2006) “representar mejor las relaciones existentes entre poblaciones mundiales, de acuerdo a su origen étnico, a nivel de marcadores microsatélites”. En los árboles obtenidos por el algoritmo UPGMA tanto para 13 y 20 marcadores se observa que la población hispanoamericana se encuentra ubicada en un grupo distinto al de la muestra poblacional mestiza peruana, esto debido a que las poblaciones hispanoamericanas están conformadas por individuos de diferente origen étnico que poseen en común la misma lengua española como lenguaje oficial (Diccionario de la lengua española, 2019); en cambio, las poblaciones caucásicas, asiáticas y africanas son agrupadas de acuerdo a su origen étnico.

Por lo expuesto se refleja que es necesario ampliar la base de datos de ADN forense peruana para que pueda representar el profundo y vasto acervo genético en nuestra población; sin embargo, este trabajo ha sido realizado con el objetivo de conocer la estructura genética actual de una muestra de residentes en Lima y/o Callao, que son originarios de diversas regiones mas no pretende generalizar los resultados para las demás poblaciones del país, ya que poblaciones estrechamente cerradas, como los Ashaninka, presentan diferenciación poblacional con la muestreada en este estudio.

En este sentido, es necesario que se realicen más estudios y con un mayor número de marcadores en otras poblaciones del país.

## 7. CONCLUSIONES

- Se detectó 8 alelos que todavía no han sido descritos en la base de datos del THE NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY (NIST). Siendo estos alelos el 12.3 del marcador CSF1PO, 8.2 del marcador TPOX, 9.1 del marcador D2S441, 11.2 del marcador D19S433, 9.2 del marcador D13S317 y 14, 19.1, 20.3 del marcador D12S391.
- Los 20 marcadores STRs en la muestra poblacional estudiada se encuentran en equilibrio de Hardy-Weinberg.
- Los 20 marcadores STRs en la muestra poblacional estudiada no se encuentran ligados.
- Los 20 marcadores STRs estudiados presentaron un Poder de Discriminación combinado de 0.999999999 y Poder de Exclusión combinado de 0.999999672, evidenciándose que los 20 marcadores en conjunto son buenos discriminantes y permiten evitar errores de falsa inclusión.
- Existe diferencias significativas entre las frecuencias alélicas de los 20 loci STRs de la muestra poblacional estudiada y la población hispanoamericana, evidenciándose subdivisión genética entre ellos.

## **8. RECOMENDACIONES**

- Se recomienda seguir realizando este tipo de estudio de distribución de frecuencias alélicas en diferentes poblaciones del país a fin de ampliar la base de datos forense de ADN que permite ser aplicada a los casos de identificación de individuos por ADN así como para pruebas de paternidad.
- Realizar este tipo de estudios con marcadores moleculares a nivel del cromosoma X e Y, así como a nivel del ADN mitocondrial, con la finalidad de ser utilizados en análisis forenses y para estudios poblacionales.

## 9. REFERENCIAS BIBLIOGRAFICAS

- Aguiar, V. R., Louro, I. D., Malta, F. S., Almeida, V. C., Ferreira, A. C., & de Castro, A. M. (2014). New CODIS core loci allele frequencies for 96,400 Brazilian individuals. *Forensic Science International: Genetics*, 1-9.
- Alleyne, L. (2009). Interpol Handbook on DNA Data Exchange and Practice: Recommendations from the Interpol DNA Monitoring Expert Group. Lyon, France: Interpol.
- Alves, C., Gusmao, L., López-Parra, A. M., Mesa, M. S., Amorim, A., & Arroyo-Pardo, E. (2005). STR allelic frequencies for an African population sample (Equatorial Guinea) using AmpFISTR Identifiler and Powerplex 16 kits. *Forensic science international*, 148, 239-242.
- Barletta, C. F. (2010). *Estudio de la estructura genética en una muestra de la población de Lima (Perú) utilizando ocho marcadores microsatélites autosómicos* (tesis de pregrado). Universidad Nacional Mayor de San Marcos, Lima, Perú.
- Botstein, D., White, R. L., Skolnick, M., & Davis, R. W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American journal of human genetics*, 32(3), 314.
- Bozzo, W. R., Lojo, M. M., Ortiz, M. I., & Pena, M. A. (2007). Genetic data from Powerplex 16 system and Identifiler kits from Buenos Aires province (Argentina). *Legal medicine*, 9, 151-153.
- Budowle, B., Chakraborty, R., & Monson, K. L. (1996). Estimating minimum allele frequencies for DNA profile frequency estimates for PCR-based loci. *International journal of legal medicine*, 108, 173-176.
- Budowle, B., Chakraborty, R., Niezgoda, S., & Shea, B. (2001). CODIS STR loci data from 41 sample populations. *Journal of Forensic Science*, 46, 453-489.
- Butler, J. M. (2005). Forensic DNA typing: biology, technology, and genetics of STR markers. USA. Elsevier.
- Butler, J. M., & Hill, C. R. (2012). Biology and genetics of new autosomal STR loci useful for forensic DNA analysis. *CRC press*.
- Butler, J. M., McCord, B. R., Crivellente, F., & Buel, E. (2004). Forensic DNA typing by capillary electrophoresis using the ABI Prism 310 and 3100 genetic analyzers for STR analysis. *Electrophoresis*, 25, 1397-1412.
- Butler, J. M., Schoske, R., Vallone, P. M., Redman, J. W., & Kline, M. C. (2003). Allele frequencies for 15 autosomal STR loci on US Caucasian, African American, and Hispanic populations. *Journal of forensic sciences*, 48(4), 908-911.

- Camacho, M. V., Figueiras, A. M., & Benito, C., (2007). Allelic frequencies of the 15 STR loci included in the AmpFISTR® Identifier™ PCR Amplification Kit in an autochthonous sample from Spain. *Forensic science international*, 173, 241-245.
- Carracedo, Á. (1996). La variabilidad genética de los micro y minisatélites y su aplicación en medicina legal. En J. Mendez (Ed.), *La genética molecular en el diagnóstico de las patologías humanas: estrategias y tecnologías* (pp. 63-74). España: Universidad de la Coruña, Servicio de Publicaciones.
- Carracedo, A., Sobrino, B., & Lareu, M. V. (2008). Forensic DNA typing technologies: a review. In *Handbook of Analytical Separations* (Vol. 6, pp. 945-957). Elsevier Science BV.
- Chakraborty, R. (1992). Sample size requirements for addressing the population genetic issues of forensic use of DNA typing. *Human biology*, 141-159.
- Chakraborty, R., & Schull, W. J. (1976). A note on the distribution of the number of exclusions to be expected in paternity testing. *American journal of human genetics*, 28(6), 619.
- Chiurillo, M. A., Ramírez, J. L., Mendes, A. M., Tovar, F., Lander, N., Fuentes, A., & Morales, A. (2003). Genetic profiling of a central Venezuelan population using 15 STR markers that may be of forensic importance. *Forensic science international*, 136, 99-101.
- De la Cruz, A. F. (2006). *Análisis de la variabilidad genética en una muestra poblacional de la provincia de Andahuaylas–Apurímac, Perú–con los marcadores microsatélites TPOX Y TH01* (tesis de pregrado). Universidad Nacional Mayor de San Marcos, Lima, Perú.
- Diccionario de la lengua española, 2019. (en línea). Disponible en: <https://dle.rae.es/?id=KW5espe>. Revisado en 1 de Junio del 2019.
- Eguiarte, L.E., Souza, V. & Aguirre, X. *Ecología Molecular* (2007). UNAM. México, D.F.
- El-Alfy, S. H., & El-Hafez, A. F. A. (2012). Paternity testing and forensic DNA typing by multiplex STR analysis using ABI PRISM 310 Genetic Analyzer. *Journal of Genetic Engineering and Biotechnology*, 10(1), 101-112.
- Evett, I. W., & Gill, P. (1991). A discussion of the robustness of methods for assessing the evidential value of DNA single locus profiles in crime investigations. *Electrophoresis*, 12(2-3), 226-230.
- Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular ecology resources*, 10, 564-567.



- Fan, H., & Chu, J.-Y. (2007). A brief review of short tandem repeat mutation. *Genomics, Proteomics & Bioinformatics*, 5(1), 7-14.
- FBI, 2018. (en línea). "Combined DNA Index System (CODIS)". Estados Unidos. Disponible en: <https://www.fbi.gov/services/laboratory/biometric-analysis/codis>. Revisado el 6 de Junio del 2018.
- Felsenstein, J. (2013). PHYLIP (Phylogeny Inference Package) version 3.695. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle, Washington.
- Fernández, F. *Análisis de 15 loci tipo short tandem repeats (STR) en la población de Paraguay para su uso en identificación forense* (tesis doctoral). Universidad de Granada.
- Filho, J. (2012). *Comparación de la frecuencia alélica de 13 loci STRs de la población brasileña y española para fines de identificación humana en genética forense* (tesis doctoral). Universidad de Valladolid. España.
- Fisher, R. A. (1951). Standard calculations for evaluating a blood-group system. *Heredity*, 5(1), 95.
- García, D. (2008). Diversidad genética y establecimiento de prioridades en esquemas de conservación: ejemplo de aplicación en la raza de lidia. Universidad Complutense de Madrid, Servicio de Publicaciones.
- Gattepaille, L. (2015). *Population Genetic Methods and Applications to Human Genomes* (Doctoral dissertation, Acta Universitatis Upsaliensis).
- Gill, P., Millican, E., Urquhart, A., Oldroyd, N., Kimpton, C. P., Sparkes, R., & Watson, S. (1996). A new method of STR interpretation using inferential logic-development of a criminal intelligence database. *International journal of legal medicine*, 109, 14-22.
- Gomes, V., Gusmão, L., Alves, C., Amorim, A., Gomes, I., Carracedo, Á., & Sánchez-Diz, P. (2009). Population data defined by 15 autosomal STR loci in Karamoja population (Uganda) using AmpF/STR Identifiler kit. *Forensic Science International: Genetics*, 3, e55-e58.
- González, J. F. (2006). *Análisis Molecular de Variación de Polimorfismos STR autosómicos y de cromosoma "Y" en grupos étnicos de Ecuador con aplicación médico-forense* (tesis doctoral). Universidad de Zaragoza, Zaragoza, España
- González-Andrade F., Martínez-Jarreta, B., & Sánchez-Q, D. (2003). Genetic profile of the Ecuadorian Mestizo population (Ecuador—South America) by using the Power Plex® 16 System Kit. *Forensic science international*, 135, 64-66.
- Goodwin, W., Hadi, S., & Linacre, A. (2011). *An introduction to forensic genetics*. Estados Unidos: John Wiley & Sons.

- Hameed, I. H., Jebor, M. A., Ommer, A. J., Yoke, C., Zaidian, H. K., Al-Saadi, A. H., & Abdulazeez, M. A. (2014). Genetic variation and DNA markers in forensic analysis. *African Journal of Biotechnology*, 13(31).
- Hameed, I. H., Mohammed, G. J., Murad, A. F., & Ommer, A. J. (2015). Allele frequency data of 21 autosomal short tandem repeat loci in Mesan and Basra provinces in South Iraq. *Egyptian Journal of Forensic Sciences*, 5, 150-156.
- Hares, D. R. (2015). Selection and implementation of expanded CODIS core loci in the United States. *Forensic Science International: Genetics*, 17, 33-34.
- Hill, C. R., Duewer, D. L., Kline, M. C., Coble, M. D., & Butler, J. M. (2013). US population data for 29 autosomal STR loci. *Forensic Science International: Genetics*, 7(3), e82-e83.
- Homburger, J. R., Moreno-Estrada, A., Gignoux, C. R., Nelson, D., Sanchez, E., Ortiz-Tello, P. & Gravel, S. (2015). Genomic insights into the ancestry and demographic history of South America. *PLoS genetics*, 11(12), e1005602.
- Huston, K. A. (1998). Statistical analysis of STR data. *Profiles DNA*, 1(3), 14-15.
- INEI, 2017. (en línea). Características Sociales de la población censada. Disponible en: [https://www.inei.gob.pe/media/MenuRecursivo/publicaciones\\_digitales/Est/Lib1539/index.html](https://www.inei.gob.pe/media/MenuRecursivo/publicaciones_digitales/Est/Lib1539/index.html). Revisado el 15 de Agosto del 2019.
- Jobling, M. A., & Gill, P. (2004). Encoded evidence: DNA in forensic analysis. *Nature Reviews Genetics*, 5(10), 739.
- Jones, D. A. (1972). Blood samples: probability of discrimination. *Journal of the Forensic Science Society*, 12(2), 355-359.
- Lewontin, R. C., & Kojima, K. I. (1960). The evolutionary dynamics of complex polymorphisms. *Evolution*, 14(4), 458-472.
- Liu, Y. S., Chen, J. G., Mei, T., Guo, Y. X., Meng, H. T., Li, J. F., ... & Zhang, L. P. (2017). Genetic variation and forensic characteristic analysis of 25 STRs of a novel fluorescence co-amplification system in Chinese Southern Shaanxi Han population. *Oncotarget*, 8(33), 55443-55452.
- Martínez-Sevilla, V. M., Aguilar-Velázquez, J. A., Inclán-Sánchez, A., Romero-Rentería, O., Díaz-Navarro, X. X., Saiz, M., & Rangel-Villalobos, H. (2016). Forensic parameters of the Powerplex® 21 system (Promega Corp.) in mestizo populations from the West of Mexico. *Spanish Journal of Legal Medicine*, 42, 10-16.
- National Geographic, 2019. (en línea). Your Regional Ancestry: Reference Populations. Disponible en: <https://genographic.nationalgeographic.com/reference-populations/>. Revisado el 15 de Agosto del 2019.

- National Research Council Committee on DNA Forensic Science (1996). The Evaluation of Forensic DNA Evidence. National Academy Press: Washington, DC; usually referred to as NRCII; recommendations listed in Appendix VI.
- Nei, M. (1987). Molecular evolutionary genetics. Columbia university press.
- Nei, M., & Roychoudhury, A. K. (1974). Sampling variances of heterozygosity and genetic distance. *Genetics*, 76(2), 379-390.
- NIST STRBase, 2019. (en línea). Disponible en: <https://strbase.nist.gov/fbicore.htm>. Revisado el 18 de Junio del 2019.
- Orellana, P. P., Andrade, C. F., Arciniegas, C. L., & Iannacone, G. C. (2017). Allele frequencies in Azuay Population in Ecuador. *Genetics and molecular research: GMR*, 16(3).
- Paredes, M., Crespillo, M., Luque, J. A., & Valverde, J. L. (2003). STR frequencies for the PowerPlex® 16 System Kit in a population from Northeast Spain. *Forensic science international*, 135(1), 75-78.
- Pereira, T. F. C., Malaghini, M., Magalhães, J. C. M., Moura-Neto, R., & Sotomaior, V. S. (2018). Genetic data for 26 autosomal STR markers from Brazilian population. *International journal of legal medicine*, 1-3.
- Pérez, L., García, O., Izarra, F., Zubiate, U., Ochoa, O., & Hau, J. (2003). Allele frequencies for the 13 CODIS STR loci in Peru. *Forensic science international*, 132, 164-165.
- Piñero, D., Salas, R., Rocha, A., Eguiarte, L., & Barahona, A (2008). La variabilidad genética de las especies: aspectos conceptuales y sus aplicaciones y perspectivas en México. En J. Sarukhàn (Ed.), *Capital Natural de México. Volumen I. Conocimiento actual de la biodiversidad* (pp 415-435). Mexico.
- Projic, P., Škaro, V., Šamija, I., Durmić-Pašić, A., Pojskić, N., Kovačević, L., Bakal, N., Primorac D., & Marjanović, D. (2007). Allele frequencies for 15 short tandem repeat loci in representative sample of Croatian population. *Croatian medical journal*, 48, 473-477
- Restrepo, T., Ibarra, A., Posada, Y., Palacio, O., Gusmão, L., Zapata, S., & Martinez, M. (2011). Database sample size effect on minimum allele frequency estimation: Database comparison analysis of samples of 4652 and 560 individuals for 22 microsatellites in Colombian population. *Forensic Science International: Genetics Supplement Series*, 3, e13-e14.
- Rey, M., Giraldo, A., Schroeder, B., Carracedo, A., Usaquén, W., & Gutiérrez, A. (2003). Allele frequencies for 13 STR's from two Colombian populations: Bogotá and Boyacá. *Forensic science international*, 136, 83-85.

- Rocabado, O., García, O., Yurrebaso, I., Inda, F. J., & Taboada, P. (2009). Population genetic data for 15 STR loci (Identifiler™ kit) in Bolivia. *Legal Medicine*, 11, 302-304.
- Ruitberg, C. M., Reeder, D. J., & Butler, J. M. (2001). STR Base: a short tandem repeat DNA database for the human identity testing community. *Nucleic acids research*, 29(1), 320-322.
- Sánchez, M. (2017). Variabilidad genética de *Attalea moorei* (Glassman) Zona "Shapaja", proveniente de cinco poblaciones naturales de la amazonia peruana. Universidad Nacional de la Amazonia peruana.
- Sandoval, J. R., Santos, F. R., Acosta, O., Fujita, R., Castillo-Herrera, W., Pena, S. D., & Salazar-Granara, A. (2013). Tracing the genomic ancestry of Peruvians reveals a major legacy of pre-Columbian ancestors. *Journal of human genetics*, 1-8.
- Semikhodskii, A., Kozub, N. A., & Sozinov, I. A. (2012). Genetic data on 15 STR loci in the Caucasian population of the Russian Federation. 53-58.
- Smith, A., & Nelson, R. J. (2004). Capillary electrophoresis of DNA. *Current protocols in molecular biology*, 2-8.
- Talledo, M., Gavilan, M., Choque, C., Aiquipa, L., Arévalo, J., & Montoya, Y. (2010). Comparative allele distribution at 16 STR loci between the Andean and coastal population from Peru. *Forensic Science International: Genetics*, 4(4), e109-e117.
- Tamaki, K., & Jeffreys, A. J. (2005). Human tandem repeat sequences in forensic DNA typing. *Legal Medicine*, 7(4), 244-250.
- Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular biology and evolution*, 30(12), 2725-2729.
- Tereba, A. (1999). Tools for analysis of population statistics. *Profiles DNA*, 2, 14-16.
- Tie, J., Wang, X., & Oxida, S. (2006). Genetic polymorphisms of 15 STR loci in a Japanese population. *Journal of forensic sciences*, 51(1), 188-189.
- Tillmar, A. (2010). Populations and Statistics in Forensic Genetics. Linköping University Electronic Press.
- Tineo, D. H. *Caracterización genética de una muestra poblacional ashaninka en el distrito de Puerto Bermúdez, Pasco-Perú, empleando marcadores STR autosómico y del cromosoma Y* (tesis de maestría). Universidad Nacional Mayor de San Marcos. Lima. Perú.

- Tito, R. Y. (2003). *Estimación de las Frecuencias Alélicas del D16S539 en una Muestra Poblacional de Ascendencia Andina Ancashina* (tesis de pregrado). Universidad Nacional Mayor de San Marcos. Lima. Perú.
- VeriFiler™ Express PCR Amplification Kit User Guide (en línea). 2017. Disponible en: [https://assets.thermofisher.com/TFS-Assets/LSG/manuals/100043588\\_VFE\\_UG.pdf](https://assets.thermofisher.com/TFS-Assets/LSG/manuals/100043588_VFE_UG.pdf). Revisado el 5 de Mayo del 2019.
- Willems, T., Gymrek, M., Highnam, G., Mittelman, D., Erlich, Y., & 1000 Genomes Project Consortium. (2014). The landscape of human STR variation. *Genome research*, 24(11), 1894-1904.
- Ziętkiewicz, E., Witt, M., Dąca, P., Żebracka-Gala, J., Goniewicz, M., Jarząb, B., & Witt, M. (2012). Current genetic methodologies in the identification of disaster victims and in forensic analysis. *Journal of applied genetics*, 53(1), 41-60.

## 10. ANEXOS Y GLOSARIO

### Anexo 1. Consentimiento informado y ficha de registro.



#### CONSENTIMIENTO INFORMADO PARA PROYECTO DE INVESTIGACIÓN CONJUNTA "COMPARACIÓN DE TECNOLOGÍAS PARA IDENTIFICACIÓN POR ADN"

- Responsables de la Investigación:** Laboratorio de Genética Humana de la UNMSM a través del grupo de investigación "Genes, Cromosomas y Genomas de Mamíferos" y la Sub Gerencia de Innovación del RENIEC
- Criterios a tener en consideración:** El presente proyecto tiene como propósito comparación de tecnologías para identificación por ADN. La persona que desee libre y voluntariamente participar en la prueba debe de ser mayor de edad y no tener vínculos de parentesco con otros participantes, la muestra tomada será sangre venosa y debido a la finalidad de la investigación no es necesario que la persona realice una dieta y/o una preparación preliminar. Este procedimiento no implica ningún tipo de riesgo. No habrá ninguna retribución económica entre las partes y las muestras serán utilizadas única y exclusivamente con fines de investigación.
- Personal de contacto:** En caso requiera mas información puede contactarse con Margarita Velasquez telefono 9870-48386 profesora Asociada del Laboratorio de Genética Humana-Facultad de Ciencias Biológicas-UNMSM y/o con Carlos Neyra teléfono 9434-98300 Biólogo de la Sub Gerencia de Innovación del RENIEC.

DATOS PERSONALES DEL PARTICIPANTE				
DNI N°:				
Nombres y apellidos				
Edad				
Correo electrónico				
Lugar de nacimiento	Distrito			
	Provincia			
	Departamento			
Antecedentes	Transferencia sanguínea	SI	NO	Fecha:
	Trasplante de médula	SI	NO	Fecha:
DATOS DE LOS PADRES		PADRE		MADRE
Nombres y apellidos				
Edad				
Lugar de nacimiento	Distrito			
	Provincia			
	Departamento			
DATOS DE LOS ABUELOS		ABUELO PATERNO	ABUELA PATERNA	ABUELO MATERNO
Lugar de nacimiento	Distrito			
	Provincia			
	Departamento			

**PARTICIPANTE:** He leído y/o se me ha leído la información contenida en esta página de consentimiento informado. He tenido la oportunidad de recibir la información necesaria respecto al proyecto y de formular preguntas relacionadas a mi participación, las cuales fueron resueltas de forma satisfactoria y entendible, además se me explicó que el estudio a realizar no implica ningún tipo de riesgo.

Por lo tanto, doy mi consentimiento voluntario para participar de este proyecto, a través de la toma de mi muestra biológica. Así mismo, tengo derecho a dejar de participar en el momento que decida, sin que ello me genere una sanción o pérdida de algún beneficio.

**PERSONAL ENCARGADO DE LA TOMA DE MUESTRA:** Confirmando que el participante ha dado su consentimiento de manera voluntaria y ha tenido la oportunidad de hacer preguntas, las cuales han sido resueltas de manera satisfactoria.

<p>_____</p> <p>Firma</p> <p>Nombres y apellidos: .....</p> <p>.....</p> <p>DNI N°: .....</p>	<p>Impresión Dactilar dedo índice derecho</p>	<p>_____</p> <p>Firma</p> <p>Nombres y apellidos: .....</p> <p>.....</p> <p>DNI N°: .....</p>
---	---	---

## Anexo 2

Constancia de Aprobación CIEI-2018-015.



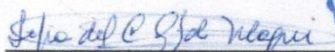
### Comité Institucional de Ética en Investigación IMT "DAC" UNMSM

#### Constancia de Aprobación CIEI-2018-015

El Comité Institucional de Ética en Investigación del Instituto de Medicina Tropical "Daniel Alcides Carrión" de la Universidad Nacional Mayor de San Marcos aprobó el 23 de octubre de 2018 la investigación "Frecuencias alélicas a partir de marcadores moleculares para identificación humana en la población peruana", a cargo del investigador principal Ricardo Jaime Descailleaux Dulanto, en los siguientes términos:

- Modalidad de revisión: revisión expedita.
- La presente aprobación del CIEI – IMT "DAC" UNMSM es por un año, del 23 de octubre de 2018 al 22 de octubre de 2019.
- El protocolo de investigación sellado por el CIEI – IMT "DAC" se encuentra adjunto a la presente constancia de aprobación. Una versión electrónica ha sido enviada también al correo de la investigadora principal: [jdescailleaux@gmail.com](mailto:jdescailleaux@gmail.com)

Lima, 23 de octubre de 2018

  
Sofía González Collantes  
Presidenta



Constancia de aprobación CIEI-2018-015  
CIEI - IMT "DAC" UNMSM – 23/10/2018

1/1

### Anexo 3

Lugar de nacimiento de los 200 voluntarios que accedieron a participar en el estudio.

Lugar de nacimiento	Voluntarios	Porcentaje
Lima	135	67.5%
Callao	17	8.5%
Junín	8	4.0%
Ayacucho	6	3.0%
Ancash	4	2.0%
Ucayali	3	1.5%
Piura	3	1.5%
La libertad	3	1.5%
Huánuco	3	1.5%
Huancavelica	3	1.5%
Tumbes	2	1.0%
Loreto	2	1.0%
Arequipa	2	1.0%
Amazonas	2	1.0%
Tacna	1	0.5%
Pasco	1	0.5%
Lambayeque	1	0.5%
Ica	1	0.5%
Cuzco	1	0.5%
Cajamarca	1	0.5%
Apurímac	1	0.5%
TOTAL	200	100%



#### **Anexo 4**

Voluntarios nacidos en Limay/o Callao cuyos ancestros pertenecen a un mismo departamento, diferente a Lima y/o Callao.

<b>Origen de procedencia</b>	<b>Voluntarios</b>
<b>Ancash</b>	6
<b>Ayacucho</b>	3
<b>Arequipa</b>	3
<b>Cuzco</b>	3
<b>Puno</b>	1
<b>Huancavelica</b>	1
<b>Cajamarca</b>	1
<b>TOTAL</b>	18

## Anexo 5

Lugar de nacimiento de los 200 voluntarios participantes en este estudio, así como el lugar de nacimiento de los padres y abuelos paternos y maternos.

Nº	LUGAR DE NACIMIENTO						
	VOLUNTARIO	PADRE	MADRE	ABUELO PATERNO	ABUELA PATERNA	ABUELO MATERNO	ABUELA MATERNA
1	Callao	Arequipa	Arequipa	Arequipa	Arequipa	Arequipa	Arequipa
2	Lima	Pasco	Pasco	Pasco	Pasco	Cusco	Cusco
3	Lima	Callao	Callao	Piura	Callao	Callao	Callao
4	Lima	Apurímac	Apurímac	Apurímac	Apurímac	Apurímac	Apurímac
5	Lima	Lambayeque	Ica	Lambayeque	Lambayeque	Ica	Ayacucho
6	Junín	Junín	Junín	Ica	Junín	Junín	Junín
7	Lima	Ayacucho	Huancavelica	Ayacucho	Ayacucho	Arequipa	Huancavelica
8	Lima	Ayacucho	Ayacucho	Ayacucho	Ayacucho	Ayacucho	Ayacucho
9	Lima	Piura	Callao	Piura	Piura	Lima	Lima
10	Lima	Lima	Huánuco	Piura	Piura	Huánuco	Huánuco
11	Cusco	Cusco	Arequipa	Cusco	Cusco	-	Arequipa
12	Lima	La Libertad	Lima	La Libertad	La Libertad	Apurímac	Lima
13	Tumbes	Pasco	Lima	Huancavelica	Cajamarca	Lima	Lima
14	Arequipa	Ancash	Puno	Ancash	Ancash	Puno	Puno
15	Tumbes	Ayacucho	Tumbes	Ayacucho	Ayacucho	-	Tumbes
16	Lima	Junín	Apurímac	Junín	Pasco	Apurímac	Apurímac
17	Lima	Lima	Lima	Tumbes	Arequipa	Arequipa	Ancash
18	Lima	Piura	Lima	Tumbes	Tumbes	Piura	Junín
19	Lima	Arequipa	La Libertad	Arequipa	Arequipa	Cajamarca	-
20	Lima	Puno	Ayacucho	-	-	Ayacucho	-
21	Lima	Lima	Lima	Lima	Lima	Lima	Lima
22	Lima	Ancash	Ancash	Ancash	Ancash	Ancash	Ancash
23	Ayacucho	Ayacucho	Ayacucho	Ayacucho	Ayacucho	Ayacucho	Cuzco
24	Callao	La Libertad	Lima	La Libertad	La Libertad	Lima	Lima
25	Lima	Callao	Lima	Callao	Lima	Callao	-
26	Lima	Lima	Lima	Lima	Lima	Piura	Piura
27	Lima	Lambayeque	Lima	Lambayeque	Lambayeque	Lambayeque	Cajamarca
28	Callao	Ayacucho	Ayacucho	Ayacucho	Ayacucho	Ayacucho	Ayacucho
29	Lima	Cusco	Cusco	Cusco	Cusco	Cusco	Cusco
30	Lima	Ayacucho	Lima	Ayacucho	Ayacucho	Piura	Apurímac
31	Lima	Lima	Lima	Ancash	Lima	-	Loreto
32	Lima	Ayacucho	Cuzco	Ayacucho	-	Cuzco	-
33	Lima	Arequipa	Ayacucho	Ayacucho	Ayacucho	Ayacucho	Ayacucho
34	Lima	Lima	Apurímac	Lima	Lima	Apurímac	Apurímac
35	Lima	Junín	Lima	Junín	Junín	Ancash	Lima
36	Lima	Ayacucho	Cusco	Ayacucho	Ayacucho	Cusco	Cusco

37	Lima	Ayacucho	Lima	Ayacucho	Lima	Lima	Lima
38	Lima	Lima	Lima	Lima	Lima	Lima	Lima
39	Lima	Lima	Lima	-	-	Lima	Ica
40	Callao	Lima	Junín	Ayacucho	La Libertad	Junín	Ica
41	Junín	Huancavelica	Junín	Huancavelica	Huancavelica	Huancavelica	Junín
42	Lima	Ancash	Cajamarca	Ancash	Ancash	Ayacucho	Cajamarca
43	Lima	Amazonas	Cajamarca	Amazonas	Amazonas	Cajamarca	Cajamarca
44	Callao	Lima	Lima	-	Lima	Callao	Callao
45	Amazonas	Amazonas	Loreto	-	Amazonas	Loreto	Loreto
46	Lima	Puno	Lima	-	Puno	-	Arequipa
47	Lima	Huánuco	Huánuco	-	-	-	-
48	Lima	Arequipa	Lima	Arequipa	Arequipa	Callao	Lima
49	La Libertad	La Libertad	La Libertad	La Libertad	La Libertad	La Libertad	La Libertad
50	Lima	Lima	Lima	Apurímac	Apurímac	-	Ayacucho
51	Lima	Junín	Lambayeque	Junín	Junín	-	La Libertad
52	Junín	Junín	Junín	-	-	-	Junín
53	Lima	Arequipa	Ayacucho	Arequipa	Arequipa	Ayacucho	Ayacucho
54	Huánuco	Huánuco	Huánuco	Huánuco	Huánuco	-	-
55	Lima	Huancavelica	Huancavelica	Huancavelica	Huancavelica	Huancavelica	Huancavelica
56	Lima	Lima	Lima	Ancash	Ancash	Ancash	Ancash
57	Lima	Ancash	Cajamarca	Ancash	Ancash	Cajamarca	Cajamarca
58	Arequipa	Arequipa	Arequipa	Puno	Cusco	Arequipa	Arequipa
59	Lima	Lima	Ancash	Puno	Arequipa	Ancash	Ancash
60	Callao	Puno	Puno	Puno	Puno	Puno	Puno
61	Lima	Cajamarca	Cajamarca	Cajamarca	Cajamarca	Cajamarca	Cajamarca
62	Lima	Lima	Ucayali	Apurímac	Apurímac	-	-
63	Lima	Piura	Lima	Piura	Piura	Lima	Lima
64	Lima	Lima	Lima	Arequipa	Arequipa	Huancavelica	Ayacucho
65	Lima	Junín	Lima	Junín	-	Ayacucho	Ayacucho
66	Lima	Huancavelica	Lima	Arequipa	Huancavelica	Huánuco	Huánuco
67	Lima	Lima	Huancavelica	Junín	Lima	Ayacucho	Ayacucho
68	Callao	Ancash	Ancash	Ancash	Ancash	Ancash	Ancash
69	Lima	Arequipa	Arequipa	Arequipa	Arequipa	Arequipa	Arequipa
70	Lima	Junín	Lima	Junín	Junín	Lima	Lima
71	Lima	Callao	Lima	Callao	Lima	San Martín	Lima
72	Lima	Lima	Callao	Ancash	Ancash	-	-
73	Lima	Lima	Junín	-	-	-	Huaraz
74	Lima	Lima	Piura	Lambayeque	Arequipa	Piura	Piura
75	Lima	-	Apurímac	-	-	Apurímac	Apurímac
76	Huánuco	Huánuco	Huánuco	Huánuco	Huánuco	Huánuco	Huánuco
77	Callao	Lima	Callao	Lima	Lima	Puno	Puno
78	Lima	Lima	Lima	Lima	La Libertad	Lima	Lima
79	Lima	Lima	Lima	Arequipa	Arequipa	Lima	Lima
80	Piura	Piura	Piura	Piura	Lambayeque	Piura	Piura
81	Lima	Arequipa	Lima	-	-	Lima	Ancash

82	Lima	Lima	Lima	Ancash	Cajamarca	Piura	Piura
83	Lima	Lima	Ancash	Lima	Lima	Ancash	Ancash
84	Huancavelica	Huancavelica	Huancavelica	Huancavelica	Huancavelica	Huancavelica	Huancavelica
85	Ayacucho	Ayacucho	Ayacucho	Ayacucho	Ayacucho	-	Ayacucho
86	Lima	Cusco	Cusco	Cusco	Cusco	Cusco	Cusco
87	Lima	Pasco	Ayacucho	Pasco	-	Ayacucho	Ayacucho
88	Lima	Arequipa	Junín	Arequipa	Arequipa	Junín	Junín
89	Lima	Lima	Huaraz	-	-	-	-
90	Lima	La Libertad	Lima	Cajamarca	La Libertad	Lima	-
91	La Libertad	La Libertad	Lambayeque	La Libertad	La Libertad	Lambayeque	Cajamarca
92	Lima	Lima	Piura	-	-	Piura	Piura
93	Lima	Apurímac	Lima	-	-	Apurímac	Apurímac
94	Huancavelica	Huancavelica	Huancavelica	Huancavelica	Huancavelica	Huancavelica	Huancavelica
95	Ayacucho	-	Ayacucho	-	-	Ayacucho	Ayacucho
96	Tacna	Arequipa	Moquegua	Arequipa	Arequipa	Arequipa	Moquegua
97	Callao	Callao	-	Callao	Callao	-	-
98	Lima	Lima	Lima	Huaraz	Lima	Lima	Lima
99	Lima	Ica	Ica	Ica	Apurímac	-	-
100	Junín	Huánuco	Huancavelica	Huánuco	Huánuco	Huancavelica	Huancavelica
101	Ayacucho	Ayacucho	Ayacucho	Ayacucho	Ayacucho	Ayacucho	Ayacucho
102	Pasco	Pasco	Apurímac	Pasco	Pasco	Apurímac	Apurímac
103	Ayacucho	Ayacucho	Ayacucho	Arequipa	Ayacucho	Ayacucho	Ayacucho
104	Loreto	Loreto	Loreto	Loreto	Loreto	Loreto	Loreto
105	Ayacucho	Ayacucho	Ayacucho	Ayacucho	Ayacucho	Ayacucho	Ayacucho
106	Lima	Lima	Lima	Lima	Lima	Ancash	Ancash
107	Lima	Ucayali	Huánuco	Ayacucho	Ucayali	Huánuco	Huánuco
108	Junín	Huancavelica	Huancavelica	Huancavelica	Huancavelica	Huancavelica	Huancavelica
109	Lima	Arequipa	Puno	Arequipa	Arequipa	Puno	Puno
110	Lima	Apurímac	Pasco	Apurímac	Apurímac	Pasco	Huánuco
111	Lima	Lima	Lima	-	Lima	Puno	Cuzco
112	Lima	Lima	Lima	Lima	-	Pasco	Lima
113	Lima	Huancavelica	Huancavelica	-	Huancavelica	Huancavelica	Huancavelica
114	Piura	Lima	Lima	Ayacucho	Huánuco	Loreto	Ancash
115	Ancash	Ancash	Ancash	Ancash	Ancash	Ancash	Ancash
116	Lima	Lima	Lima	Apurímac	Ancash	La Libertad	Ancash
117	Lima	Junín	Lima	Huancavelica	Junín	Huancavelica	Huancavelica
118	Lima	Lima	Lima	Ancash	Lima	Ancash	Ancash
119	Lima	Junín	Piura	Lambayeque	Junín	Huánuco	Piura
120	Ucayali	Lima	Ucayali	Loreto	Ica	Huánuco	Ucayali
121	Huánuco	Ayacucho	Arequipa	Ayacucho	Ayacucho	Arequipa	Arequipa
122	Lima	Lima	Lima	Lima	Lima	Ayacucho	Ayacucho
123	Lima	Huancavelica	Lima	-	-	Arequipa	Loreto
124	Ucayali	Lima	Ucayali	Lima	Ancash	Ucayali	Ucayali
125	Amazonas	Amazonas	Amazonas	Amazonas	Amazonas	Amazonas	Amazonas
126	Lima	Ancash	Ayacucho	Cajamarca	Ancash	Ayacucho	Ayacucho

127	Callao	Ancash	Ancash	Ancash	Ancash	Ancash	Ancash
128	Junín	Junín	Junín	Pasco	Junín	Junín	Junín
129	Lima	Lima	Lima	Ancash	Ancash	Pasco	Ancash
130	Lima	Ancash	Ancash	Ancash	Ancash	Ancash	Ancash
131	Lima	Ayacucho	Lima	Ayacucho	Ayacucho	Cajamarca	Lima
132	Lima	Lambayeque	Lima	Lambayeque	Lambayeque	Lima	Huánuco
133	Lima	Ancash	Ancash	Ancash	Ancash	Ancash	Ancash
134	Lima	Lima	Piura	Ica	Lima	Piura	-
135	Lima	Lima	Lima	Lima	Junín	Ancash	Ancash
136	Apurímac	Apurímac	Apurímac	Apurímac	Apurímac	Apurímac	Apurímac
137	Lima	Lima	Lima	Ancash	Ancash	Ancash	Ancash
138	Junín	Pasco	Lima	Pasco	Pasco	Lima	Ayacucho
139	Lima	Junín	Junín	Huancavelica	Junín	Junín	Junín
140	Callao	Callao	San Martín	Piura	Piura	San Martín	San Martín
141	Ica	Ayacucho	Ica	Cajamarca	Ica	Arequipa	Ica
142	Callao	Callao	San Martín	Puno	Puno	San Martín	San Martín
143	Lima	Cajamarca	Lima	Cajamarca	Cajamarca	Ayacucho	Huancavelica
144	Lima	Lima	Ayacucho	-	-	Ayacucho	Ayacucho
145	Lima	Huancavelica	Lima	Huancavelica	Huancavelica	Pasco	Lambayeque
146	Ucayali	Amazonas	Ucayali	Amazonas	Loreto	Loreto	Loreto
147	Lima	Cajamarca	Amazonas	Cajamarca	Cajamarca	Amazonas	Amazonas
148	Lima	Ayacucho	Ancash	Ayacucho	Ayacucho	Ancash	Ancash
149	Lima	Lima	Ancash	Lima	Ayacucho	La Libertad	Ancash
150	Lima	Cusco	Huancavelica	Cusco	Cusco	Junín	Junín
151	Lima	Lima	Lima	Lima	Cusco	Lima	Lima
152	Callao	Lima	Lima	Amazonas	Ica	Ayacucho	Ayacucho
153	Lima	Ancash	Lima	Ancash	Ancash	Lima	Lima
154	Lima	Lima	Lima	Apurímac	Lima	Ica	Junín
155	Huancavelica	Ica	Huancavelica	-	-	Lima	Huancavelica
156	Lima	Piura	Lima	Piura	-	-	-
157	Lima	Lima	Lima	Ica	Lima	Lima	Lima
158	Callao	Lima	Lima	Ica	Lima	La Libertad	Lima
159	Lima	La Libertad	Lima	La Libertad	La Libertad	Lima	Cusco
160	Lima	Lima	Lima	Piura	Lima	Lima	Lima
161	Lima	Arequipa	Arequipa	Arequipa	Arequipa	Arequipa	Arequipa
162	Ancash	Ancash	Ancash	Ancash	Ancash	Ancash	Ancash
163	Lima	Lima	Lima	Arequipa	Arequipa	Junín	Junín
164	Lima	Lima	Cusco	Lima	Lima	Cusco	Cusco
165	Lima	Ayacucho	Ayacucho	Ayacucho	Ayacucho	Ayacucho	Ayacucho
166	Lima	Cusco	Cusco	Cusco	Cusco	Cusco	Cusco
167	Piura	Piura	Piura	Piura	Piura	Piura	Ucayali
168	Lima	Apurímac	Lima	Apurímac	Apurímac	Arequipa	Arequipa
169	Lima	Huancavelica	Lima	-	-	-	-
170	Lima	Lima	Lima	Lima	Lima	Lima	Lima
171	Lima	Lima	Lima	La Libertad	Piura	Cajamarca	Huancavelica

172	Lima	Lima	Ucayali	Junín	Junín	Ucayali	Ucayali
173	Lima	Junín	Lima	La Libertad	La Libertad	Junín	Lima
174	Callao	Lima	Lima	Lima	Lima	Ancash	Piura
175	Lima	Callao	Lima	Ancash	Ancash	Ayacucho	Ancash
176	Lima	Cajamarca	Ica	Cajamarca	Cajamarca	Lima	Ica
177	Ancash	Ancash	Ancash	Ancash	Ancash	Ancash	Ancash
178	Lima	Lima	Apurímac	Lima	Lima	Apurímac	Apurímac
179	Lima	Lima	Lima	Huancavelica	Ancash	Apurímac	Apurímac
180	Cajamarca	Cajamarca	Cajamarca	Cajamarca	Cajamarca	Cajamarca	Cajamarca
181	Lima	Ancash	Apurímac	Ancash	Ancash	Apurímac	Apurímac
182	Lima	Lima	Lima	Ayacucho	Junín	Huánuco	Ica
183	Lima	Junín	-	Junín	Junín	-	-
184	Lima	Lima	Lima	Lambayeque	Lima	Ayacucho	Lima
185	Callao	Lima	-	Lima	-	-	-
186	Lima	Lima	Lima	Lambayeque	Ayacucho	Ayacucho	Ayacucho
187	Lima	Ayacucho	Ayacucho	Ayacucho	Ayacucho	Ayacucho	Ayacucho
188	Lima	Ayacucho	Huánuco	Ayacucho	Ayacucho	Huánuco	Huánuco
189	Loreto	San Martín	Loreto	San Martín	San Martín	Loreto	Loreto
190	Callao	-	Callao	-	-	Cajamarca	-
191	Lambayeque	Lambayeque	Lambayeque	Lambayeque	Lambayeque	Lambayeque	Lambayeque
192	Junín	Junín	Junín	Junín	Junín	Junín	Junín
193	Lima	Lima	-	Apurímac	Apurímac	-	-
194	Lima	Lima	Lima	Piura	Piura	Lima	Lima
195	Lima	Ancash	Ancash	Ancash	Ancash	Ancash	Ancash
196	Lima	Lima	Lima	Lima	Lima	Piura	Lima
197	La Libertad	La Libertad	Cajamarca	Cajamarca	La Libertad	Cajamarca	Cajamarca
198	Ancash	Lima	Ancash	-	-	Ancash	Ancash
199	Lima	Puno	Ayacucho	Arequipa	Puno	Ayacucho	Ayacucho
200	Lima	Lima	Junín	Ica	Ica	Junín	Junín